TABLE OF CONTENTS

CHAPTER 1 5

HISTORICAL ANTECEDENTS TO EI 5
THE READY ACCEPTANCE OF EI 6
WHAT EXACTLY IS EI? 8
EI AND ITS PSYCHOMETRIC CORRELATES 10
ABILITY EI AS A PREDICTOR 11
SELF-REPORT EI AS A PREDICTOR 12
EI AS A UNIFIED CONSTRUCT 12
THE BUSINESS CASE FOR EI 15
OVERVIEW OF THE THESIS 16
CLOSING INTRODUCTORY REMARKS 18

CHAPTER 2 19

THE TWO TYPES OF EI 20
SUMMARY OF AIMS AND HYPOTHESES 23

METHOD 23

TASKS IN THIS BATTERY OF TESTS 23
ETHICS 30
APPARATUS 30
TEST PROCEDURE 30

RESULTS 31

DESCRIPTIVE STATISTICS 31
TEST BATTERY 32
GROUP DIFFERENCES 32
GENDER DIFFERENCES 33
OUTLIERS 33
CORRELATION MATRIX 34
H1: EI TESTS WOULD BE SIGNIFICANT PREDICTORS OF THE BEB CRITERION 36
REGRESSION ANALYSIS 36
HIERARCHICAL REGRESSION ANALYSIS 38
H2: THE MAGNITUDE OF THE VARIANCE EXPLAINED WOULD JUSTIFY FURTHER RESEARCH 41
H3: THE TESTS WOULD CLUSTER INTO THEORETICALLY DEFINED GROUPS 41
CONFIRMATORY FACTOR ANALYSIS 43

DISCUSSION 46

INCREMENTAL VARIANCE EXPLAINED BY EI 46
CONSTRUCT VALIDATION 47
EVALUATING THE TESTS 48
LIMITATIONS AND FUTURE DIRECTIONS 49
CONCLUSIONS 50

CHAPTER 3 51

EI AS A UNIFIED INTELLIGENCE 52
EI TESTS USED IN THIS STUDY 53
SUMMARY OF AIMS AND HYPOTHESES 54

METHOD 55

PARTICIPANTS 55
BIOGRAPHICAL QUESTIONNAIRE 55
TASKS IN THIS BATTERY OF TESTS 55
EI TESTS USED IN THIS STUDY 59
ETHICS 62
APPARATUS 62

Test Procedure 62
RESULTS 62
Descriptive Statistics 62
Behavioural observations 63
Correlation Matrices 64
Testing Hypothesis 1 67
Testing Hypothesis 2 71
Testing Hypothesis 3 74
Factor analysis for this study 76
DISCUSSION 79
EI as a single factor (H1) 79
EI as a predictor of maturity and superior interpersonal skills (H2) 80
EI as a predictor of first-semester psychology scores (H3) 83
Replicating the Ford and Tisak study 84
Limitations and Strengths of this Study 85
CHAPTER 4 86
The Bandwidth-Fidelity dilemma 88
Choosing tests for this study 88
Summary of aims and hypotheses 90
METHOD 90
Tasks in this Battery of Tests 91
Biographical Details 92
Emotional Intelligence tests 92
Big-Five Factor Personality Inventory (BFF) 93
General intelligence markers (IQ) 94
Miscellaneous inventories 94
Ranking of performance of Sales Representatives 96
Test Procedure 97
RESULTS 98
Descriptive Statistics 98
Correlation Matrix 99
Relationship between EQ-i subfactors and the criterion 100
H1: Total EI scores will predict the Performance Ranking 101
H2: Total EI scores will be better predictors than IQ scores 101
H3: Selected EI subfactor scores will be significant predictors of performance 101
H4: Selected EI and miscellaneous variables will be significant predictors 103
H5: Tests that predict the difference between Average and Excellent performers 103
DISCUSSION 104
Total EI scores as predictors of performance (H1) 104
Total EI scores will be better predictors than IQ scores (H2) 105
Selected EI subfactor scores as significant predictors of performance (H3) 106
Miscellaneous test scores will add incremental variance (H4) 108
Certain tests will identify superior performers (H5) 109
Limitations of this Study 109
Summary 110
CHAPTER 5 111
Summary of aims and hypotheses 114
METHOD 115
Tasks in this Battery of Tests 115
Biographical Details 116
Emotional Intelligence tests 116
Big-Five Factor Personality Inventory (BFF) 117
General intelligence markers (IQ) 118
Miscellaneous inventories 118
Manager report – The criterion 119

Apparatus 120
Test Procedure 120

RESULTS 120

Descriptive Statistics 120
Skew in the criterion (PKPI) 121
Correlation Matrix 123
Relationship between the Performance KPI criterion and EI subscales 125
H1: Total EI scores as better predictors of the PKPI criterion than IQ 125
H2: EI subscales will explain more variance than total EI scores 126
Stepwise regression analysis 128
Tests that predict the difference between average and excellent performers 130
H3: Total EI scores will discriminate between average and excellent performers 130
H4: Selected tests will explain significant variance in the PKPI criterion 131

DISCUSSION 132

General observations 132
Total EI scores as predictors of performance (H1) 132
EI as a unified intelligence (H2) 133
Total EI scores as discriminators of performers (H3) 135
Selected tests as a model (H4) 135
Limitations of this study 136
Summary 136

CHAPTER 6 138

Overall aims and findings of this thesis 138
EI as a unified construct 141
EI as a superior predictor to IQ 143
Employing EI for staff selection purposes 144
Alternative explanations and potential criticisms 144
Issues for future research 145
Implications and concluding remarks 147
Epilogue 147

SUMMARY 149

REFERENCES 151

Chapter 1

TIME magazine successfully captured the spirit of an era in October, 1995, when it featured Emotional Intelligence (EI) as the cover story. EI had been popularised by Daniel Goleman[1] in his seminal publication "*Emotional Intelligence: Why it can matter more than IQ*" (Goleman, 1995). Goleman's book found a ready audience following Herrnstein and Murray's (1994) controversial tome, reviewed one year earlier in TIME that emphasised the primacy of IQ in determining life outcomes. Herrnstein and Murray asserted that IQ was primarily genetically determined and unchangeable; that the difference between black versus white IQ scores were persistent, and that low IQ was implicated in many of society's problems ranging from poverty, illegitimate births, to predicting jail populations. In contrast, Goleman (1995) optimistically asserted that EI mattered more than IQ in achieving life success, could be raised by EI developmental interventions in school and the workplace, and was a more appropriate measure for today's service-oriented society. This was a contrasting and uplifting message for those who believed they had suffered from comprehensive testing in schools and been placed in an IQ category for which there seemed no escape. Since 1995, the ascendancy of EI over IQ has continued as EI has been widely assimilated into the workplace and interpersonal lexicon (Matthews, Zeidner & Roberts, 2002). Although Goleman popularised the EI concept, Mayer and Salovey had defined the term and had been developing and delineating EI from IQ since 1990 (Salovey & Mayer, 1990).

Historical antecedents to EI

While Intelligence Quotient (IQ) testing had been used successfully for selection in the schools and recruiting in World War I, it was argued by many that abstract and verbal reasoning tasks did not provide a complete picture of intelligence; especially performance with regard to interpersonal situations (Mackintosh, 1998). Thorndike proposed that the definition of intelligence be broadened beyond the narrow constraint of academic intelligence with the addition of the concept of Social Intelligence (SI), using the following definition: "Social intelligence is the ability to understand and manage men and women, boys and girls - to act wisely in human relations." (Thorndike, 1920, p. 228). However, apart from a notable study by Ford and Tisak (1983) (discussed in the second and third chapters of this thesis) researchers were unable to isolate an SI factor.

In the middle of last century David Wechsler, acknowledging Thorndike's SI contribution, put forward the view that there are additional non-intellective capacities such as: conscientiousness,

[1] While Goleman is an oft-cited authority on EI due to his many popular publications, the author notes that this does not constitute peer-reviewed scientific literature. However, since Goleman's publications have set the agenda that has largely been taken up, often without critical debate, by the general community such as education and organizational psychology, it is necessary to reference and test some of his claims.

personal interest, emotional factors, impulsivity, and curiosity that go beyond logical and abstract reasoning; these he proposed should be included under the banner of intellectual ability (Wechsler, 1940, 1950). While he subsequently included some aspects of social intelligence such as picture arrangement into his tests, he generally did not follow through with the issues he raised in 1950. By 1958, Wechsler's position on SI had moved to the point where he stated that, "social intelligence is just general intelligence applied to social situations" (Wechsler, 1958, p. 75).

In the early 1980s, Howard Gardner introduced his theory of multiple intelligences which included interpersonal and intrapersonal intelligence (Gardner, 1983, 1999). Although some of his "intelligences" such as body kinaesthetic intelligence came under criticism from the psychometric establishment for simply being abilities rather than intelligences per se, Gardner's theories were widely adopted in the educational realm, where his more egalitarian rather than the previous elitist doctrine of intelligence was more useful. Due to the diversity of the multiple "intelligences," Gardner's theory does not support a common intelligence factor such as psychometric *g*. The concept of multiple intelligences, whereby every child could find their own special niche, added a social and educational dimension to the previous narrow conception of intelligence. It is probable that the germ of the idea for Goleman's first EI book was born when Goleman and Gardner were both postgraduate students studying at Harvard in the late 1970s (e.g., Goleman, 1996).

The term "Emotional Intelligence" first came into use in the early 1990s (Salovey & Mayer, 1990); although there are references to the use of the term before that date (e.g., Payne, 1986). However, it was not until 1995, after the publication of Goleman's first book, that the term Emotional Intelligence (EI), or Emotional Quotient (EQ), reached worldwide consciousness; due in no small part to being featured in TIME magazine (Gibbs, 1995). From that point on, EI attracted the attention of a sizeable part of the psychological research community who in the ensuing 12 years have attempted to validate the EI construct (e.g., Davies, Stankov & Roberts, 1998, Petrides & Furnham, 2001, 2000a; Brackett & Mayer, 2003; Ciarrochi, Forgas, & Mayer, 2001; Dulewicz & Higgs, 2000), and attain the high levels of predictive performance alluded to in Goleman's book. As a measure of the uptake and interest in EI, there were four references for EI listed in PsycINFO in 1995, rising to 199 articles listed in 2006, with a grand total of 1136 EI references between 1995 and 2007.

The ready acceptance of EI

It is interesting to ponder why, at the end of the 20th century, the major focus in individual differences shifted from rational-cognitive processes to the importance of emotional and interpersonal adaptation. The rational philosophy that arose during the enlightenment that dominated most of the 20th century stressed that humans, as higher-level beings, had to forgo their lower-level emotional nature and aspire to be influenced only by rational-cognitive processes

(Izard, 2001). This advice dates back to Aristotle who advised that the wise man should express appropriate levels of emotion in situations that aroused anger and passion (from *Ethics*, cited in Goleman, 1995). The Bell Curve (Herrnstein & Murray, 1994) appeared to be the last straw of cognitive elitism in a politically correct new age, and a willing general public enthusiastically seized the new, egalitarian concept of EI.

Two contexts that immediately recognised the potential of EI were schools and organizations. Schools have found it increasingly challenging to adapt to the expectation that they address moral and social issues, in comparison to earlier times where authoritarian schoolmasters simply administered the curriculum under the rule of corporeal punishment. A more homogeneous and religiously-oriented community to that which exists today freed the curriculum teachers from involvement in moral and ethical issues, leaving the development of moral values and virtues to the church and brief religious education classes provided by visiting ministers (e.g., Fagan, 2006). In the latter part of the 20th century, single-parent homes, latch-key students, drugs, alcohol, teen pregnancies, and violence in schools have compounded the duties of the contemporary teacher to complete the basic academic curriculum. Teachers have struggled to complete the academic curriculum, while being confronted by the social problems listed above on a daily basis. The 60s movie "*To Sir with love*" presented an early insight into the needs of teenage students that went beyond the basic academic curriculum (Braithwaite & Clavell, 1967). In the 90s, emotional-literacy courses were introduced to address many of the social problems faced by students. By introducing soft-skills topics into the classroom that address feelings of exclusion, bullying, needs for affection, and awareness of the emotions of others, educators have begun to address a variety of social problems, potentially enabling the student to better attend to the academic curriculum (Goleman, 1995). However, this thesis does not address the application of EI in high schools; rather, it explores the impact of EI on success in first-year university, and success in organizations.

In the West, where profit, continuous improvement, and the belief in working smarter, not harder is the general approach to increased productivity, organizations have long sought tools that could predict superior performance in the workplace. Hiring superior candidates means that performance benefits are realised earlier, and there is not the expensive and wasteful process of attrition (e.g., Spencer, 2001). Goleman proposed that in terms of success in business, EI was a superior predictor to IQ and could account for up to 80% of the "forces" that make up success, while IQ could at best only explain 20% of the contributing factors (Goleman, 1995, p. 34). Up to that point, IQ had been considered the pre-eminent predictor of success at work; built on the belief that academic success leads to success at work. However, IQ has demonstrated limited success in predicting superior employee performance (e.g., Sternberg & Hedlund, 2000); and when taken to the limit, IQ only predicts potential success for an academic career (Sternberg, 2000a).

Although research into EI as a predictor of success has escalated since 1995, this expansion has often faltered due to varying definitions of EI, a lack of transparency in publications, such as not providing the data for further analysis, and a limited number of peer-review of studies in the organizational environment (Matthews et al., 2002).

What exactly is EI?

While extolling the importance of EI and its advantages over IQ, little in terms of theory and definitions of EI were included in Goleman's first book (Goleman, 1995). There were some broad associations of EI with *character*, *self-discipline*, and leading a virtuous life, but the reader had to wait until his next book to find a clearer definition of the components of EI (Goleman, 1998). The assertion by Goleman that EI was everything not covered by general intelligence (IQ) was too broad a definition (Hedlund & Sternberg, 2000). The theory underlying EI appeared to be a work-in-progress as there were no validated EI tests in the mid 1990s (Goleman, 2001a). An early sample EI quiz published by Goleman (1995) did some disservice to the EI concept as it turned out to have major reliability problems (see Davies, Stankov & Roberts, 1998). Goleman (1998) responded with an emotional competence framework for EI consisting of a mix of 25 subfactors, from which grew the Emotional Competence Inventory (ECI) (Boyatzis, Goleman & Rhee, 2000).

Meanwhile, Salovey and Mayer had been working on defining and delineating EI from IQ from as early as 1990 (Salovey & Mayer, 1990). These scholars have continued to defend EI as an intelligence based on Wechsler's broad definition that "intelligence is the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with the environment" (Wechsler, 1939, p. 3). This led to their current definition of EI as: "the ability to (1) accurately perceive emotions, (2) to use emotions to facilitate thinking, problem solving , and creativity; (3) understand emotions; and (4) manage emotions to for personal growth" (Mayer, Salovey & Caruso, 2002c, p. 1). This definition forms the basis for the latest (v.2) Mayer, Salovey and Caruso Emotional Intelligence test (MSCEIT); an ability-based scale where participants identify emotions in faces pictures, and solve emotion-oriented problems and situations (Mayer, Salovey & Caruso, 2002c, 2002d). The questionnaire is veridically scored and participants receive a score for each item based on consensus with a large sample of the previous test-takers in the database. The MSCEIT is defined as an "ability" or "performance test", with right and wrong answers, in contrast with self-report tests such as the EQ-i that ask the participant to rate their skills subjectively on a Likert scale.

Reuven Bar-On[2], another major proponent of EI who authored the EQ-i test, defines EI as: "a self-report measure of emotionally and socially competent behaviour that provides an estimate of one's emotional and social intelligence" (Bar-On, 2000, p. 364). The EQ-i arose from Bar-On's (1980) studies that attempted to ascertain why some people exhibit better psychological adjustment to life than others (Bar-On, 1997a). The EQ-i invites participants to:

> "describe yourself by indicating the degree to which each statement is true of the way you feel, think or act most of the time and in most situations" on a 1 to 5 scale where: 1 represents *not true of me*, and 5 represents *very often true of me* (Bar-On, 1997, p. 1).

This instrument is a "self-report" test where there are no right or wrong answers and the participant rates their own feelings, views and style as they see fit. This type of EI test has also been designated a "mixed" EI test as it contains a mix of 15 subscales addressing a wide range of abilities ranging from Social responsibility, to Happiness (Roberts, Zeidner & Matthews, 2001; Forgas, Mayer & Ciarrochi, 2001). Petrides and Furnham (2001) proposed that performance EI tests such as the MSCEIT should be designated as "Information-processing" or "Ability" EI tests, and self-report tests such as the EQ-i and the TEIQue should be designated "Trait" EI.

One of the major shortcomings of EI is that the ability and trait EI tests do not appear to measure the same construct, as shown in a study where the MSCEIT and the EQ-i only correlated at 0.21, indicating a shared variance between the tests of only 4% (Brackett & Mayer, 2003). While this would be expected to raise doubts and threaten to bring disrepute to the whole concept of EI, the EI advocates appeared to ignore this issue and continued to extol the virtues of EI without addressing the root problem that there are now two or more different definitions of the EI construct.

With the two versions of EI, self-report and ability, continuing to gather momentum in educational and organizational contexts, there appears to have been no major attempts merge the two theories into a unified EI definition. Mayer et al. continue to define EI as an intelligence as measured by an array of items with predetermined answer scores derived from consensus scoring, while Bar-On and other developers of self-report EI inventories do not appear to be making any move towards modifying their theories to include performance EI items. In terms of the currently available EI tests, the MSCEIT (Mayer, Salovey & Caruso, 2002c) stands alone as the only reputable performance EI test. The self-report or trait EI tests are more numerous, for example: the EQ-i (Bar-On, 1997), the Assessing Emotions Scale (AES) (Schutte et al., 1998), the Trait Emotional Intelligence Questionnaire (TEIQue) (Petrides, Perez & Furnham, 2003), the Emotional

[2] Bar-On is well-known for his popular EQ-i test and publications, rather than through the peer-reviewed, scientific literature channels. However, his importance as a theorist cannot be underestimated as the EQ-i test is a widely-used and oft-cited instrument. It is in this context that Bar-On is being cited.

Competency Inventory (ECI) (Boyatzis, Goleman & Rhee, 2006), the Emotional Intelligence Appraisal (EIA) (Bradberry & Greaves, 2004), and the SUEIT (Palmer & Stough, 2001).

There are other factors that complicate the issue of merging definitions of EI, such as Sternberg's assertion that *practical* and *tacit knowledge* address the same domain as EI (Sternberg, 2000, 2000a; Hedlund & Sternberg, 2000), and Epstein's (1998) exhortation that *constructive thinking* as measured by his own CTI instrument, is the key to EI. In summary, the difficulties involved in merging the multiple definitions of trait and ability EI appear to be insurmountable at this stage.

EI and its psychometric correlates

An issue that exacerbates the differences between the two theories is that ability EI is expected to be significantly positively correlated with general intelligence (IQ), and divergent from the FFM of personality; however, self-report EI is typically found to be correlated with personality, and divergent from IQ scores (Matthews, Zeidner & Roberts, 2002). Mayer et al. refer to the correlation of EI (as measured by the MSCEIT) with IQ scores as evidence that EI is an "intelligence" per se (Mayer et al., 1993, 2001). One of the defining characteristics of a general intelligence battery that dates back to Spearman (1904) is that subtests must demonstrate positive correlations with each another; referred to as the "positive manifold" (Mackintosh, 1998). The positive manifold has been proposed as evidence for a common general (*g*) or IQ factor behind individual general cognitive intelligence subfactors, such as verbal and abstract reasoning. Thurstone, on the other hand, argued that the positive manifold simply represents common method variance such as pencil and paper test-taking skills, reading and comprehension, and not a general intelligence factor (Thurstone, 1941, cited in Mackintosh, 1998). However, the positive manifold concept has generally become the major defining characteristic for a general intelligence factor, and the major justification for the inclusion of new intelligence subfactors (e.g., Carroll, 1993; Guttman & Levy, 1991).

When Mayer et al. found that the MSCEIT demonstrated a positive manifold within its subfactors and with general intelligence tests, this was presented as evidence that EI was indeed an intelligence, and not simply a competency or skill (Mayer et al., 2001). However, Thurstone's criticism could be applied when examining a "blends" item from the MSCEIT that asks: "Another word for 'consistently anticipating pleasure' is? *a. optimism*, *b. happiness*, *c. contentment*, *d. joy*, *e. surprise*" (Mayer Salovey & Caruso, 2002d. p. 23). This item looks very similar to items found in verbal IQ tests; it requires common reading and comprehension skills, and it comes as no surprise that this item representing the "blends" subfactor of EI is correlated with general intelligence. Following Thurstone's reasoning, the positive manifold exhibited between EI subfactors may be

due in part to common method variance; thus, insufficient evidence for subfactors such as "blends" being categorised as a new intelligence per se.

The positive manifold between subtests of the MSCEIT is necessary, but not sufficient, evidence for EI being a new intelligence. As far back as 1936 Allport performed a study of trait names and noted approximately 18,000 descriptive terms (Allport, 1936, 1958). While many of these terms were found to be related to each other, forming constructs, and exhibiting positive correlations with other constructs, this was insufficient evidence for grouping them and labelling them a new intelligence. The International Personality Item Pool (IPIP) lists 269 scales, further demonstrating that there is no shortage of constructs divergent from IQ and personality (Goldberg, 1999). Similarly, of the many divergent constructs possible, such as the MSCEIT subfactors, some of which demonstrate a positive manifold, there is insufficient evidence to call them an intelligence until it can be shown what criterion they actually predict. In the case of the MSCEIT, while there have been modest positive correlations between total EIQ scores, and workplace and educational criteria, Mayer et al. have not attempted to specify the criterion that the MSCEIT actually predicts. The self-report EI tests also predict workplace and educational outcomes to some extent revealing that the MSCEIT appears to have no unique predictive advantage that can be conferred upon it due to it being an ability measure, and putative status as an intelligence (e.g., Mayer et al., 1999).

Ability EI as a predictor

At the dawn of IQ testing, Binet (1905) designed tests to identify the criterion of students in need of remedial education (Mackintosh, 1998). In a chapter entitled "What does EI predict?" Mayer et al. define three criteria that the MSCEIT predicts: empathy ($r = .33$), parental warmth ($r = .23$), and life satisfaction ($r = .19$) (Mayer, Caruso & Salovey, 2000). While this is a start, there are there issues that present themselves: the correlations are small; and, the 138 item MSCEIT test may be a case of overkill. It is possible that a 10 to 20 item test of empathy, for example, would do the same and probably a better job. Finally, the criteria are *post hoc* findings as the MSCEIT was never designed to predict these criteria. The MSCEIT thus appears to be a solution in search of a problem, or simply the results of research along the lines of, "If you don't know where you are going, any road will get you there." (Carroll, 1865).

On the other hand, the self-report EI inventories such as the EQ-i are not without their limitations and have often been criticised as being too highly correlated with the Big Five Factor (BFF) or Five Factor Model (FFM) of personality (Costa & McCrae, 1985, 1992). Although this criticism is persistent and applies to all self-report EI tests (e.g., Matthews, Zeidner & Roberts, 2002; Petrides & Furnham, 2001, 2004), the MSCEIT also correlates significantly with factors in the FFM. For example, the MSCEIT correlates significantly ($p < .001$) with two factors in the FFM, while the EQ-i correlates significantly with four of the FFM factors (Brackett & Mayer, 2003).

These results indicate that correlation with the FFM of personality is more an issue of degree than a problem that is inherent only in self-report EI scales.

Self-report EI as a predictor

The EQ-i documentation states that it was designed to address the issue of, "why are some individuals more able to succeed in life than others?" (Bar-On, 1997, p. 1). This criterion provides researchers and general consumers of the EQ-i with some guide as to where it may be usefully applied. Knowing the criterion that the EQ-i was designed to address suggests areas such as: success at work, in marriage, education, and overall self-actualization in life, as suitable applications areas for the EQ-i.

EI as measured by the EQ-i does not purport to be a new intelligence per se as Goleman (1995) intimated that EI was in his seminal publication, or as Mayer et al. (2001, 2002c) claim for the MEIS and the MSCEIT. The EQ-i claims are more conservatively stated; EI is "part of general intelligence and can be measured along a continuum" (Bar-On, 1997, p. 147). The issue becomes one of whether a new inventory addressing success in life and psychological well-being needs to be defined as an intelligence to be useful and worthy of further research. Establishing EI with the credentials of a new intelligence (Mayer, 1993, 1999, 2001; Goleman, 1995; Matthews, Zeidner & Roberts, 1992), as distinct from a family of loosely-related social skills (Ciarrochi & Godsell, 2006), has consumed time and valuable journal space, but has not shed any new light on real-world application of the EI concept.

In summary, it appears that arguments regarding EI's convergence with general intelligence (e.g., Mayer et al., 1999, 2000, 2001, 2002c), and significant correlations with the FFM of personality indicating poor divergent validity (Matthews et al., 2002), have been somewhat overstated and academic, and have held little interest for the general consumers of the MSCEIT and EQ-i tests. While it would be hoped that superior divergent validity from the FFM could be developed in due course, the situation regarding convergent validity between EI tests is less optimistic; the EQ-i and the MSCEIT displaying poor convergent validity with each other with only 4% of variance overlap (Brackett & Mayer, 2003). For EI to achieve psychometric validity there is an urgent need to establish convergence with real-world criteria such as success at work and interpersonal relationships, and to determine the effect that EI has on different job roles and vocations. In this regard some initial studies in the area of "EQ profiling" for particular occupations has commenced using the EQ-i (Bar-On, Handley & Fund, 2006a).

EI as a unified construct

Central to the promise that a total or global EI score can be used to predict real-world criteria (e.g., Goleman, 1995, 1998) is the requirement that there is a positive manifold between

tests and that EI subfactors contribute to a unified structure with the global EI score as the unitary high-level factor. Taking the general intelligence construct as a model with general intelligence (*g*) at the top as a template (e.g., Carroll, 1993), there must be sufficient common variance within subfactors to render the total EI score meaningful as a global measure. The claim that there is a single global EI score that satisfactorily represents the eight subfactors in the MSCEIT or the 15 subfactors in the EQ-i is supported to some extent by exploratory and confirmatory factor analysis (CFA) (e.g. Bar-On, 1997, 2000; Mayer et al., 1999, 2002c).

However, the evidence for a unified structure often does not stand up to scrutiny, for example, as the EQ-i only shows a CFA model of the total EQ score and the five composite scores, without including the any of the 15 subfactors (Bar-On, 1997). Confirmatory Factor Analysis (CFA) demonstrated support for only 10 out of the 15 EQ-i factorial components, that is, a 10-factor structure instead of the theoretical 15-factor structure (Bar-On, 2000). This indicates that CFA cannot confirm a unified structure for the EQ-i model, and raises doubts regarding the ability of the total EQ score to adequately represent the 15 underlying factorial components.

The question arises as to whether a construct such as EI must exhibit a unified structure to be considered a useful addition to the existing catalogue of psychometrically valid constructs. This issue appears not to have concerned the consumers of these tests to date as they continue to use total EQ scores in the same manner as IQ scores. However, it would be valuable if researchers addressed the unified structure issue and explained the implications to the general users of these tests.

The practice of adding diverse factorial subscales together to produce a total score, as is currently done with EI tests, is analogous to adding apples and oranges; the total score has no clear meaning as to what it represents (e.g., Michell, 1990). With 15 diverse subfactors in the EQ-i, ranging from independence to reality testing, it is difficult to visualise what composite outcome factor the summing of all these scores represents. Summing the subfactors to produce a total EQ score can only make sense when supported by empirical evidence such as CFA, and this has not been satisfactorily demonstrated to date. To sum diverse factors without this evidence dilutes the overall predictive properties of the subfactors, rendering the total EI score as less predictive than the individual subfactors.

If the requirement that subfactors exhibit a unified construct supporting a meaningful total EI score was removed, EI scales such as the EQ-i could still be useful, but the concept of a unified EI factor that is similar in application to an IQ score would no longer be tenable. Goleman's (1995) promise that EI means more than IQ would have to be reviewed or redefined because there would be no meaningful total or global EI score to reference. However, the lack of a meaningful global EI score may not restrict the applicability of EI in the long run, as models that do not provide a single

global score such as the FFM of personality, the MMPI and the 16-PF are still be very useful and predictive instruments.

The FFM model of personality (Costa & McCrae, 1985, 1992) is a valuable construct even though it does not generate a global score. The MMPI (Hathaway, McKinley, Butcher, Dahlstrom, Graham & Tellegen, 1989) is a multi-factorial inventory, with no global score, that has wide applicability in social and vocational applications. The 16-PF (Cattell, 1994) is an inventory consisting of a broad range of 16 factors, without a global score; the 16-PF has been found to be useful in a wide range of psychological applications. By relinquishing the somewhat artificially-created global EI score until it could be established empirically, EI research could continue to flourish. It is noteworthy that the two leading EI tests, the MSCEIT and the EQ-i were developed between five and 10 years ago, and little progress in terms of availability of new tests has been made in this popular area of EI research in the ensuing years.

Moving away from the emphasis on global EI scores that are expected to explain more variance and be more predictive than IQ, a recent study appears to heading in that direction (Bar-On, Handley & Fund, 2006). In this study, five examples of predicting high performers in recruiting and military units showed that EQ-i scores were able to explain variance ranging from 24% to 30% in the outcome criterion. This supports Goleman's (1995) assertion that EI matters more than IQ that typically only explains up to 20% of the variance. On closer examination, however, the researchers appear to have selectively chosen the EI subfactors that were significant in the regression analysis for each scenario. In study one, predicting the skills of USAF recruiters, seven out of the 15 EQ-i factors were selected to explain 28% of the variance. The second study, predicting successful IDF recruits, found four out of the 15 EQ-i factors, each different to the factors in study one, predicted 30% of the variance in outcomes. Study three included four factors, while studies four and five used five and nine out of the 15 available factors respectively. While there is no intrinsic reason why researchers cannot drop many or most of the available EQ-i subfactors, post hoc, after the results are evaluated, this method did not discuss the total EQ score, leading the reader to conclude that the total EQ score did not explain significant variance.

This approach (viz., Bar-On et al., 2006) yields substantial variance in the criterion through stepwise regression models; an automatic statistical process that selects the significant EQ-i subfactors for a particular sample. These results can clearly demonstrate that EI matters more than IQ in terms of overall variance explained, and also provides a useful model for selecting new recruits. However, this approach raises issues regarding capitalising on chance that some of the 15 EQ-i predictors will turn out to be significant, and yet the same predictors may not replicate in another sample (Babyak, 2004). Bar-on et al. indicate that EI is entering an "EQ profiling" phase; developing specific models that will predict vocational performance thus, "improving human

effectiveness and organizational productivity" (Bar-On, Handley & Fund, 2006, p. 17). Clearly, the concept of the one-size-fits-all global EI score originally conceived by Goleman and others appears to be no longer relevant in the "EQ profiling" phase. One would hope that EI theory is revised and disseminated widely demonstrating true scientific transparency. However, the implications of not using the total EQ score, and selectively choosing subfactors, will probably pass unnoticed by general users of EI tests, and the message simplified to being further evidence that EI is more predictive than IQ.

The business case for EI

There are two categories of utility and advantages that can accrue from a wider application of EI and theory; the soft and hard benefits. The soft benefits include the promise of a happier, more polite and considerate society, and reduced incidents of bullying and adaptation problems at school (e.g., Goleman, 1995; Saklofske & Parker, 2005; Mayer & Cobb, 2000). EI has also been found to be a significant predictor of success in high school and university (e.g., Schutte et al., 1998; Petrides, Frederickson & Furnham, 2004; O'Connor & Little, 2003; Brackett & Mayer, 2003; Goetz, Frenzel, Pekrun & Hall, 2005). One of the studies in chapter 3 addresses EI as a predictor of academic achievement; however, the main focus of this thesis is on predicting maturity, social competence, and the "hard" or commercial benefit of predicting performance in the workplace.

While some research has shown that EI predicts positive outcomes in the workplace (e.g., Goleman, 1995, 1998, 2001a, 2001b, 2002; Cherniss, 2000, 2001, 2004), these studies are typically proprietary, and do not publish the data for peer researchers to verify claims, or explore "what-if" scenarios (Matthews Zeidner & Roberts, 2004). In fact, the bar for EI adding incremental variance as a predictor for a real-world criterion with significant "consequences" has been set as low as 4% (Brackett & Mayer, 2003, p. 1156). The low incremental variance and unspecified predictive properties of EI, plus the lack of a common theoretical model are major causes for concern for the validity of the construct (e.g., Ciarrochi, Forgas & Mayer, 2001; Matthews et al., 2002, 2004).

Notwithstanding the major problems in EI measurement instruments, the promoters of EI continue to extol its global applicability and advantages to consumers. Cherniss (2001) proposes that EI relates to organizational effectiveness in a wide range of areas including: employee commitment, morale, health and retention, development of talent, teamwork, innovation, productivity, efficiency, sales, revenues, service quality, and customer loyalty; Dulewicz and Higgs (1998) propose that resilience, influence, assertiveness, integrity, and leadership are EI abilities (cited in Goleman, 2001a). Goleman's Emotional Competence Framework includes: self-awareness, accurate self-assessment, self-confidence, self-control, trustworthiness, conscientiousness, adaptability, innovation, achievement drive, commitment, initiative, optimism, understanding others, developing others, service orientation, leveraging diversity, political awareness, influence,

communication, conflict management, leadership, change catalyst, building bonds, collaboration, and teamwork (Goleman, 1998, p. 26). The majority of these claims are made without empirical studies to support them, and when studies are cited they are typically unpublished, short on detail (see Matthews et al., 2004), and unavailable to peer-researchers (e.g., Cherniss, 2004). Where peer-reviewed research has been conducted, the manifestations of EI are somewhat narrower (Matthews et al., 2004).

The application of EI in workplace situations for selecting superior employees who will increase productivity and profit, and reduce the expense of turnover and attrition, is considered to be a "hard" benefit of EI. The EI literature abounds with examples of EI scores being positively related to benefits in business performance; a few of which are cited below: consultants scoring high on EI competencies delivering 139% incremental gain; insurance agents high in EI selling policies of twice the value of average performers; and, the primary cause for substandard performance in executives being due to low EI skills (Cherniss, 2004; Goleman 1995, 1998; Cherniss & Goleman, 2001a; Caruso, Salovey & Mayer, 2003). The issues with the studies cited are that a very broad range of definitions of EI are utilised, statistical data are not published to enable peer-review, and in some cases, the studies use unpublished EI tests.

Overview of the thesis

To avoid the problem of a variety of definitions of EI, this thesis utilises EI tests that are widely available to the research community such as the AES (Schutte et al., 1998), the TMMS (Salovey, Mayer, Goldman, Turvey & Palfai, 1995), the TEIQue (Petrides & Furnham, 2004), and recognised standard EI tests such as the EQ-i (Bar-On, 1997) and the MSCEIT (Mayer et al., 2002c). Where the author has used unpublished scales, these are included in the appendix.

The overall aim of the thesis was to examine effectiveness of EI as a predictor of performance in a range of situations where there was a valid criterion for EI scores to be statistically compared with. The four studies in this thesis will address seven questions relating to EI that urgently need to be answered. They are: (1) do EI test results predict a mature behavioural style? (2) do EI scores predict academic performance? (3) what are some of the different properties between ability and self-report EI? (4) does EI explain all incremental variance over IQ and personality? (5) is EI an unified intelligence that can be adequately represented by a common factor? (6) is EI a superior predictor to IQ? and, (7) can EI be usefully employed for staff selection purposes?

Chapter 2 examines the hypothesis that maturity and good character, as displayed in a behavioural interview, are positively related to EI (e.g., Goleman, 1995). The veracity of this claim is assessed with a variety of putative self-report and ability EI tests; the results of which are correlated with a behavioural interview score. Chapter 2 also provides the data to determine the magnitude of the variance that EI subtests explain in the criterion, and to decide whether this

magnitude is sufficient to provide justification for further research. The data from the study described in this chapter are also used to explore convergence and divergence from personality and general intelligence (IQ) measures.

Chapter 3 describes a study similar to that described in chapter 2, but with improved methodology and the inclusion of up-to-date EI tests, in an attempt to provide superior prediction of the criterion. The overall focus of this chapter is to explore the predictive properties of EI and to determine the proportion of incremental variance explained by EI, personality and IQ tests. Compared with the first study, this study provides more rigorous control over the behavioural evaluation of maturity through the use of videotaped interviews, and four judges, to improve the reliability of the behavioural assessment score.

Chapter 3 also addresses the relationship between EI and academic achievement using first-semester Psychology 104 results as the criterion. EI has been found to predict success in the first year at university to varying extents (e.g., Schutte et al., 1998; Petrides, Frederickson & Furnham, 2004; Brackett & Mayer, 2003). The relative contributions of measures of IQ, EI and FFM personality have considerable implications for university entrance selection, continuing success, and remediation of students with difficulties. The incremental validity of EI over the FFM personality scores and IQ markers is examined in relation to the criterion of academic results achieved.

Chapter 4 describes a study of EI in an organizational setting where the ability of EI scores to predict high-performing sales representatives is put to the test. It has been claimed that EI is a necessary skill for success and that sales representatives high in EI typically outperform average sales representatives by achieving double the revenue (Spencer, 2001). This study examines the efficacy of EI as a predictor of sales performance in relation to the criterion of a ranking of relative performance supplied by the sales manager. The contribution of personality and IQ scores are compared and contrasted with EI scores, and the relative contribution of the ability of each to explain the ranking criterion is estimated. From the wide battery of psychometric tests administered in this survey, the tests that distinguish between high and low performers are identified.

Chapter 5 assesses the ability of EI to predict superior performers in a call-centre environment. In the area of customer insurance claims and loss inquiries, it is expected that call-centre employees with a superior customer-service orientation would be better equipped to complete the claims transactions successfully (e.g., Spencer, 2001). In this regard, skills such as empathy, emotional acuity and managing one's own emotions maps directly between EI theory and customer-service orientation skills. In this study a battery of EI, personality, IQ and miscellaneous tests are administered, and results compared with an objective performance criterion for each

employee supplied by the manager. This provides the data to determine the relative ability of each category of test to explain variance in the criterion.

This study also provides an opportunity to determine whether EI exhibits the structure of a unified intelligence. The claim that EI is a unified intelligence with a common general factor has been instrumental in the rapid rise in acceptance of EI as a new concept (e.g., Goleman, 1995; Mayer et al., 1999). However, it is also possible that EI is simply a collection of individual factors such as contained in the 16-PF, the MMPI, or the FFM of personality that may have a preferred or socially-correct pole, but do not contain sufficient common variance to support a global EI factor. The study described in chapter 5 provides an opportunity to gain some insight into whether EI is a unified intelligence or a loosely-related group of constructs (e.g., Ciarrochi & Godsell, 2006).

Finally, tests that distinguish between high and low performers are identified in Chapter 5. After determining the subtests that predict success in a specific work environment, these subtests may be administered to prospective candidates with the resulting scores indicating the extent to which they would be expected to be successful in the new job. The determination and use of predictive tests is similar to the methodology employed by Bar-On et al. with "EQ profiling" (Bar-On, Handley & Fund, 2006, p. 17).

Closing introductory remarks

In summary, the studies in this thesis attempt to empirically examine the predictive properties and the robustness of the EI construct, with the general aim of confirming or refuting much of the hyperbole surrounding EI (e.g., Goleman, 1995, 1996, 1998, 2001a, 2001b, 2002, 2006; Cherniss, 2000, 2001, 2001a, 2004). The widespread application of the EI concept, representing a focus on inter and intra-personal issues (e.g., Gardner, 1983, 1999; Mayer et al., 1993, 2001; Goleman, 1995, 1998), has added to the richness of discussion relating to educational, organizational, personal relationships and leadership issues. Rather than focusing on the shortcomings of EI on technical issues (e.g., Matthews et al., 2004) it would be preferable to retain the new and valuable insights that EI provides, and rationalise some of the mythology regarding the omnipotent nature of EI as a predictor of performance and happiness (e.g., Goleman, 1995). It is in this positive spirit of enquiry that the following four studies of EI are presented to the reader for critical evaluation.

Chapter 2

How does Emotional Intelligence (EI) manifest itself in the real world? Does it comprise good character and maturity (Goleman, 1995, 1998), superior intrapersonal and interpersonal skills (Gardner, 1999), or wisdom and the ability to manage others (Thorndike, 1920)? The term EI has been used to encompass a wide variety of constructs; for example, Constructive Thinking has been proposed as the foundation of EI (Epstein, 1998). And if it were possible to settle on a concrete definition of the real-world, behavioural manifestations of EI, what vocations and roles would high EI be valuable for?

Much has been written about Emotional Intelligence (EI) and the advantages accrued by emotionally intelligent individuals (e.g., Goleman, 1995; Caruso, 2001, Cherniss, 2000, 2001). There are a variety of psychological organizations administering EI tests and enhancement programs, typically based on a limited set of proprietary offerings, usually the MSCEIT or the EQ-i. However, these two leading EI tests, the MSCEIT (Mayer, Salovey & Caruso, 2002), and the EQ-i (Bar-On, 1997) do not measure the same construct, as shown by results from 200 participants whose global EI scores only correlated at 0.21, thus indicating a shared variance of only 4 % between the two EI tests (Brackett & Mayer, 2003). This correlation is regarded as small (Matthews, Zeidner & Roberts, 2002; Mayer et al., 2002c; Cohen, 1988), and results would be interpreted as statistically insignificant and evidence for divergent constructs for a sample size of less than 90. Although Brackett's larger sample size (N = 200), rendered this low correlation significant, it is borderline; not compelling evidence that these two EI tests are measuring the same thing.

The current study was designed to begin to address these questions by establishing the construct and predictive validity of Emotional Intelligence (EI) as described by its key proponents (e.g., Goleman, 1995; Salovey & Mayer, 1990, Bar-On, 1997). The study was designed to investigate how a wide variety of emotion-related tests such as emotion recognition in faces and voices (e.g., Ekman, 2002) and putative EI tests relate to observed maturity and socially-adept behavioural characteristics in participants.

A diverse battery of psychological tests relating to emotion recognition, standard intelligence, personality, and EI was administered to participants in a single session. At the end of the survey, a brief evaluation of behaviour (BEB) was conducted for each participant. The aim of the study was to establish which tests in the battery were significantly related with the social skill rating obtained from the interview, the brief evaluation of behaviour (BEB). This study was envisaged as a preliminary step in determining EI's relationship with superior performance in a variety of situations that demand social skills.

The two types of EI

Researchers have attempted to explain the anomalies between EI tests by proposing that there are "trait EI" or "mixed" tests using self-reports such as EQ-i, and there are "performance" or "ability" EI tests such as the MSCEIT (Petrides & Furnham, 2001; Roberts, Zeidner & Matthews, 2001; Matthews et al., 2002). Unfortunately, this very important distinction has received limited attention from organizations administering EI tests in the field, and generally does not get passed on to consumers. Psychological testing organizations are typically only authorised to administer one type of EI test, and this becomes the dominant paradigm (e.g., Kuhn, 1962; Goleman, 2001a) for the interpretation and discussion of EI with their clients. If the two major EI tests (EQ-i and the MSCEIT) share a common variance of only 4 %, it is clear that clients are receiving only a fraction of the story when their staff are tested and provided with feedback on the results of one test only. A more specific approach to EI, with clear definition of the behavioural attributes that an individual with high EI scores is expected to exhibit in real life, needs to be developed by researchers and presented to the academic community and consumers of these tests.

Understanding the differences between self-report and performance EI tests is necessary to explore the issues regarding the different EI constructs that they appear to be measuring. Performance EI tests, such as employed in the MSCEIT (Mayer et al., 2002c), test for maximum performance, typically have a single correct answer, and are an independent measurement of the individual's ability, not subject to exaggeration or bias (Matthews et al., 2002; Hedlund & Sternberg, 2000). Self-report tests, often referred to as trait EI or mixed EI tests such as the EQ-i (Bar-on, 1997), ask the test-taker to evaluate their own emotional abilities on multi-point Likert scales. Although self-report tests are easier to administer, they are subject to bias, exaggeration and demand characteristics affecting the veracity of the responses. Performance EI tests are found to correlate with general intelligence measures, while self-report tests correlate with personality measures such as the Big Five (BFF) (Matthews et al., 2004).

During the early debates (e.g., Davies, Stankov & Roberts, 1998), regarding the validity of the EI construct alluded to in the original Goleman (1995) EI book, developers were content to consider divergence from personality and IQ tests as evidence for the unique psychological construct of EI (e.g., Mayer, Caruso & Salovey, 1999). However, divergence in itself is of no value without some indication of what the EI construct predicts. While ability EI and Trait EI (Petrides & Furnham, 2001) exhibit some divergence from established constructs such as IQ and personality (e.g., Bar-On, 1997), they also exhibit divergence from each other; typically demonstrating a low correlation of 0.21 between test results (Brackett & Mayer, 2003). Bar-On claims a "relatively low" correlation of 0.22 is sufficient evidence for divergent validity (e.g., Bar-On, 1997, p. 128), and

Mayer states that constructs that correlate at less than 0.25 are considered, "unrelated to minimally related" (Mayer et al., 2002c, p. 38). Thus, since the ability and self-report EI test results are essentially unrelated, the behavioural manifestations of EI are expected to be quite different for the two types of tests, and it is currently unclear as to what the different tests predict. If the operational definition of EI is simply left at "what the tests test" (e.g., Boring, 1923), this does not provide a solid grounding for future EI research; neither does it provide consumers of these tests a formal definition of what behaviours to expect in high-scorers in the two different types of tests.

The major promoters of the EI concept (e.g., Cherniss, 2001; Emmerling & Goleman, 2006; MHS, 2006; Hein, 2006; Goleman, 2001a; Mayer et al., 1999, 2002c; Bar-On, 1997) continue to support and promote activities that line up behind the "cheerleading" role of EI (Matthews, Roberts, & Zeidner, 2004, p. 192) without due diligence in demonstrating what their products actually predict. By contrast, personality tests (e.g., Costa & McCrae, 1995; McCrae & John, 1992) clearly define the behavioural manifestation of their tests with factors such as Extraversion. Intelligence tests (e.g., Wechsler, 1997) clearly state that their tests predict superior academic performance in a variety of verbal and performance situations. This presents cause for concern for EI, considering that over 10-years of research into EI has been conducted since Goleman's (1995) publication, and still the major EI tests are essentially divergent; overlapping with a negligible 4% shared variance.

In a study by Schutte et al. (1998) using the Assessing Emotions Scale (AES), significant differences were found in the scores of therapists versus prisoners; the AES predicted superior grade point averages (GPA) of a sample of students; females scored higher in the AES than males; and the AES was insignificantly correlated with student aptitude tests of academic intelligence (Schutte, Malouf, Hall, Haggerty, Cooper, Golden & Dornheim, 1998). Brackett found that the EQ-i predicted drug use, alcohol use, and social deviance, while the MSCEIT predicted social deviance, high school rank and college GPA (Brackett & Mayer, 2003). Petrides et al. reported that the Bar-On EQ-i predicted the following criteria: emotional control (ECQ), satisfaction with life, and rational coping styles. In a second study they found that the self-report Trait Emotional Intelligence test (TEIQue) predicted criteria such as: depression, dysfunctional attitudes (DAS), and rational coping styles as measured by self-report tests (Petrides, Perez-Gonzalez, & Furnham, 2007). The criteria are derived from self-report tests and are in essence the results of comparing self-report EI tests with self-report criteria, as distinct from observable and verifiable actions that would provide more grounded real-world measures of behaviour. Nonetheless, this peer-reviewed research provides some insight into the sort of outcomes that have been scientifically related to EI, as a distinct contrast to the expansive list of attributes put forward by the promoters of EI.

Thus, the expansive list of personal characteristics related to EI as proposed by the EI-promoters (e.g., Goleman, 1998; Cherniss, 2001) has not been validated by scientific

peer-reviewed research. On the other hand, the criteria that have been validated by research, such as prisoners versus therapists, GPA scores, depression and coping styles (e.g., Schutte et al., 1998; Brackett & Mayer, 2003; Petrides et al., 2007), are too clinical and narrow for workplace applications. Furthermore, the dominant EI instruments, the EQ-i and the MSCEIT, have minimal overlap, and do not predict the same criteria, as demonstrated by the Brackett and Mayer study (2003).

Since no clear definition of EI-influenced behaviours applicable to the workplace has emerged from EI studies to date, research must continue if EI is to find a place in the current taxonomy of formal psychometric constructs, such as personality and IQ. This study was therefore designed as a first step in the process to position and establish EI as a formal construct.

Goleman (1995) proposed that, "There is an old-fashioned word for this growth in emotional intelligence: *maturity*" (1998, p. 7) and that, "emotional intelligence represents: *character*" (p. 285). This study uses these general definitions of EI proposed by Goleman to rate behavioural characteristics exhibited by participants in an interview, using a methodology similar to that employed by Ford and Tisak (1983).

In the search for a social intelligence (SI) factor, Ford and Tisak (1983) managed to isolate a "distinct social intelligence factor" (p. 196); a unique outcome that had not been demonstrated since Thorndike had first proposed the existence of a social intelligence factor (Thorndike, 1920). Ford and Tisak administered a battery of social competency, personality and academic intelligence measures to participants, followed by a behavioural interview to assess each participant's social intelligence. Factor analysis subsequently demonstrated a unique social SI factor that was divergent from personality and academic scores.

Following the Ford and Tisak methodology, this study will examine how mature behaviour, measured in an interview situation, relates to a variety of EI scales and a variety of psychometric tests. The behavioural EI score obtained from the brief evaluation of behaviour (BEB) demonstrated in the interview will become the criterion against which all putative tests of EI will be evaluated. The rationale for the study was that if EI tests correlated positively with the behavioural observation of maturity and good character in individuals, this would provide confirmatory evidence for the unsubstantiated assertions offered by Goleman (1995, 1998). It would also furnish experimental evidence for the capability of EI tests to predict the behavioural outcome of *maturity* and good *character*. Moreover, the actual magnitude of the variance explained would indicate whether EI tests have sufficient predictive validity to be utilised in workplace applications. In this regard, Mayer et al. have proposed that constructs that explain an additional 5% of the variance beyond existing psychological constructs are worthy of further research (Mayer, Salovey & Caruso, 2000b).

Summary of aims and hypotheses

The overall aim of this study was to replicate the Ford and Tisak (1983) study, and to establish whether putative tests of EI, or any other tests, could predict a behavioural outcome that represents *maturity* and good *character* (BEB). The overall aims can be summarised in three formal hypotheses:

H1: That EI test scores would be significant predictors of the BEB criterion, above and beyond the established measures of personality and general intelligence.

H2: That the magnitude of variance explained by EI tests would provide justification for further research.

H3: That EI tests would cluster into ability and self-report groupings as proposed by EI theorists, and would exhibit divergence from personality and general intelligence measures.

Method

One hundred and three first-year psychology students (78 female, 75%), from the Department of Psychology, Macquarie University, took part in this study as part of their course requirements. Their ages ranged from 18 to 41 years with a mean of 22.25 years (SD=4.17). The ethnic background of 70 (68%) of the participants was Australian or Western European, leaving 33 (32%) from non-English speaking backgrounds, predominantly Asian. The first language of participants corresponded closely with country of birth, with 72 (70%) of the participants reporting English as their first language. This mix reflects the Sydney metropolitan demographic with a population of 72% from Western and 28% from non-English speaking backgrounds (ABS, 2001).

Tasks in this Battery of Tests

A wide variety of tests designed to cover a range of psychological domains were administered to establish convergent and divergent validity with Emotional Intelligence tests and observed behaviour. A summary of the experimental battery is listed below in Table 1.

Table 1

Battery of tasks performed by each participant in this study.

Task	Type	Author/Source
1. Biographical Details (Bio)	Biographical	-
Personality Inventory		
2. BFF Personality (OCEANIC)	BFF	Roberts (1999)
General Intelligence tasks		
3. Word Reasoning Task (Word)	Gc	Stankov & Roberts (2001)
4. Matrix Reasoning (Matrix)	Gf	Stankov & Roberts (2001)
5. Trail Making Task (ZVT)	Gf	Oswald (1987)
Emotional Intelligence and related scales		
6. Tacit learning task (Tacit)	SLT	Nissen & Bullemer (1987)
7. Facial Affect Recognition (Faces)	Emotion	Mazurski & Bond (1993)
8. Brief Affect Recognition Task (BART)	Emotion	Ekman (METT) (2003)
9. Vocal Affect Recognition (Voices)	Emotion	Scherer (2002)
10. Situational Judgement Task (SJT)	Emotion	Reid (2004)
11. Emotional Stroop Task (STROOP)	Emotion	Sharma & McKenna (2001)
12. Assessing Emotions scale (AES)	EI	Schutte (1998)
13. Trait Meta-Mood Scale (TMMS)	EI	Salovey & Mayer (1995)
14. Brief Evaluation of Behaviour (BEB)	Interview	Ford & Tisak (1983)

To enable the identification of the emotional intelligence factor, the test battery included multiple tests purporting to measure the same psychological construct, as is commonly performed with factor analytic methods or multitrait-multimethod theory (Campbell & Fiske, 1959). Six performance EI scales were included in the battery, as the ability to solve emotional problems is considered by some theorists to be a more valid measure of EI than self-report inventories (e.g., Mayer, Salovey & Caruso, 2000b). The six performance EI tasks were: tacit learning; facial affect recognition; brief affect recognition; vocal affect recognition; situational judgement; and the emotional stroop task. These six tasks sample a wider domain of performance EI than the MSCEIT, and it was expected that at least some of them would produce a robust performance EI factor.

Three tests, word-reasoning, matrices, and trail making were included as markers for general intelligence. The OCEANIC (Roberts, 1999), based on Big-Five factor (BFF) theory (McRae & John, 1992) was used as a measure of personality. The Sequential Learning Task (SLT), reputed to measure tacit knowledge learning ability (Nissen & Bullemer, 1987; Berry, 1993) was included to test whether tacit knowledge is the basis of EI as proposed by Sternberg (e.g., Hedlund & Sternberg, 2000).

Biographical details. The biographical questionnaire was designed to collect the following information:

1. Date of birth (optional)
2. First language
3. Country of birth
4. Universities Admission Index (UAI)
5. Semester I psychology I mark
6. Study major if not Psychology

The Universities Admission Index (UAI) is a measure of the student's overall achievement in the final year (12) at high school when compared with all other students. The UAI is considered the best single predictor of future success in tertiary studies. The UAI and the psychology I mark were included to potentially provide additional academic intelligence criteria.

Personality Inventory (BFF). This 60-item scale is based upon items and theory underlying the "Big-Five Factor" (BFF) model (Costa, 1985; McCrae, 1992) and provides a measure of the five global factors, Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (OCEAN). This scale was developed at Sydney University and is referred to as the OCEANIC test (Roberts, 2000). The response format is a six point Likert-type scale with six quantifiers, *never, rarely, sometimes, often, usually, and always*. An example of an item from each factor is as follows:

Openness:	I have thought about the origins of the universe
Conscientiousness:	I would describe myself as efficient.
Extraversion:	I like to be where there is lots of action
Agreeableness:	I try to be kind to everyone I know
Neuroticism:	I get upset easily

For brevity, these five factors will be referred to by the one-letter abbreviations of O, C, E, A, N respectively throughout this thesis.

Tacit learning task (Tacit). This test was included in the battery as it has been suggested that tacit knowledge or practical intelligence may be the underlying factor determining emotional intelligence (Hedlund & Sternberg, 2000). This task, a putative test of tacit knowledge (Berry & Dienes, 1993; Hunt, 2001), tests the participant's reaction time to pressing keys on the computer keyboard in response to a stimulus that appears on the screen in one of four positions. The stimulus was a red spot, one centimetre in diameter, that appeared in one of four positions spaced six centimetres apart, linearly, on the screen. This design was based on Nissen and Bullemer serial reaction time (SRT) task (1987).

When a specific stimulus appears in one of the four positions, the participant must press a corresponding key on the keyboard. The requisite keys are shown at the bottom of the screen and z, x, period, and / were chosen as they provide a comfortable resting position for fingers corresponding to position 1 to 4 respectively. The test has three phases: 10-practice trials, 40-random sequence trials, 96 repeated-sequence trials, and finally, 40 random-sequence trials. The repeated sequence consists of eight digits corresponding to positions (13243142) repeated 12 times.

Tacit learning occurs during the 96 repeated-sequence trials and is measured as the difference between average random-sequence response time and the average repeated-sequence response time.

Facial affect recognition task (Faces). This task utilised the Mazurski and Bond (1993) pictures of facial affect, supplied with standardised rating information. The participants were asked to choose the emotion in each face using a subset of 20 of the best face stimuli that included male, female, adult and child (Mazurski & Bond, 1993). The six emotions displayed in a balanced and random order were: *anger, disgust, fear, happiness, sadness,* and *surprise*. Ratings were collected from participants using continuous analogue sliders that could be slid into position, thus providing a quantitative rating resolution to five decimal places. The test collected 120 ratings per participant overall (6 emotions x 20 faces) that can be analysed in a variety of ways.

Brief Affect Recognition Task (BART). This instrument was included to evaluate facial emotion recognition ability of a brief stimulus (Micro emotion training task, METT, Ekman, 2003, 2003a). The METT utilises the same stimuli as the Japanese and Caucasian brief affect recognition task (JACBART) (Matsumoto, LeRoux, Wilson-Cohn et al., 2000). The test comprises 56 stimuli from Ekman's METT showing a brief, one-fifteenth of a second emotion, forward and backward-masked by the same face without expression. After each stimulus is presented, the participant is asked to click a button to rate the expression as one of the following seven emotions: *Happiness, Contempt, Disgust, Sadness, Anger, Surprise, or Fear*.

Vocal Affect Recognition Task (Voices). This task is based on Scherer's vocal emotion recognition test (K. R. Scherer, personal communication, February 13, 2003; Scherer, Banse, & Walbott, 2001). The participants were presented with a series of 30 sound bites of vocal emotion spoken by male and female German radio actors. A synthetic language was used to eliminate semantic content that could help the participant determine the emotion expressed. The actors deliver one of two phrases that are a mixture of German, English, French, Italian, Spanish, and Danish. The two phrases are:

"Hat sundig pron you venzy."

"Fee gott laish jonkill gosterr."

The participant's task is to identify the emotion expressed by the voice by clicking the button that best describes the emotion. Choices are: *Anger, Fear, Joy, Sadness* and *Neutral.*

Situational judgement task (SJT). The Situational Judgement Task is modelled on similar tasks used in EI tests such as the MSCEIT (Mayer, Salovey & Caruso, 2002c) that claim to provide a performance measure of EI rather than simply a self-report of emotional management ability (e.g., Petrides & Furnham, 2000, 2000a, 2001, 2003). The participant was presented with a problem scenario and then presented with six possible responses to the issue raised in the scenario. The participants are then asked to rate each response using a slider scale to indicate the effectiveness of each option to resolve the issue.

Example:

> Susan's boss asked her to make a presentation to an important client, giving her only two days notice to prepare. She had never presented before and felt very nervous about speaking in public. She worried that she might jeopardize future business from the client if her presentation does not go well. The following would be likely to reduce her anxiety in the long term.

Responses for the participant to rate in terms of effectiveness:

a) Rehearsing by presenting the paper in front of a friend
b) Requesting that a colleague attend with her to the presentation
c) Telling herself that she would do her best but if it didn't go well, it wasn't her fault
d) Soldiering on, suppressing her feelings of anxiety
e) Mentally redefining the project as an opportunity for development and promotion
f) Practise her yoga relaxation exercises

The emotional Stroop task (Stroop). The Stroop task is based on the theory that attentional bias causes delays in processing under certain circumstances (e.g., Isenberg, 1999; Sharma & McKenna, 2001). In the emotional Stroop test it is postulated that emotional words will be processed at a different rate to neutral words. Participants were asked to name the font colour of a mixture of emotional and neutral words, randomised and presented singly. The emotional Stroop effect is measured by the average response time difference between emotional and neutral words. Stimuli and methodology followed the Sharma and McKenna study (2001) with maximum time pressure created by zero inter-stimulus delay between the words. Examples of stimulus words are:

Neutral: NOTE, CLOCK, THUMB, FIELD, ROSE

Emotional: FAIL, FEAR, CRASH, GRIEF, DEATH

Word reasoning task (gc). The 23 items comprising this task are taken from the Gf-Gc battery builder developed at Sydney University (Stankov, 2000). It includes items of semantic

relationships such as: “STATUE is to SHAPE as SONG is to …”. Optional answers are: “BEAUTY, PIANO, TUNE, NOTE.” The participant was asked to choose the word that most closely completes the given analogy. This task serves as a marker for the broad ability of crystallised intelligence (gc) (Stankov, 1997; Stankov & Roberts, 2001).

Matrices (gf). The Matrices Test is similar to Raven’s Progressive Matrices (Raven, 1938) and contains 10 items that become progressively more difficult, each with a one minute time limit. The participant is presented with nine large boxes each containing what looks like a "noughts and crosses" or tic-tac-toe board. Eight of the nine boxes contain a progression that goes from left to right, top to bottom. The task is to select the correct box from the five options provided to complete the progression in the ninth box. The matrix task is designed to tap the broad ability of fluid intelligence (gf) (Stankov & Roberts, 2001).

Assessing Emotions Scale (AES). The AES, a putative measure of self-report EI is based on the model proposed by Salovey and Mayer (Salovey & Mayer, 1990). The test has been made freely available for research by the authors (Schutte et al., 1998). The AES consists of 33 items with a five-point Likert scale: 1 = Strongly Disagree, to 5 = Strongly Agree.

Examples of the AES scale are as follows:

(1) I know when to speak about my personal problems to others

(2) When I am faced with obstacles I remember times I faced similar obstacles and overcame them

(3) I expect that I will do well in most things I try

Trait meta-mood scale (TMMS). The TMMS was the first EI test produced by Salovey and Mayer (Salovey, Mayer, Goldman, Turvey & Palfai, 1995; Mayer & Stevens, 1994). The short-form of the TMMS scale consists of 30-items with a five-point Likert scale with end points of: 1 = *Strongly Disagree*, and, 5 = *Strongly Agree*. The scale is based on the Salovey and Mayer EI theory (Mayer & Salovey, 1993), and includes a factorial structure consisting of Attention, Repair, and Clarity dimensions of emotions. Examples of items are as follows:

Attention: People would be better off if they felt less and thought more (r)

Repair: I try to think good thoughts no matter how bad I feel

Clarity: Sometimes I can’t tell what my feelings are (r)

Trail Making Task (Tracing). The trail-making test, also known as the Tracing Task, is based on the Der Zahlen-Verbindungs test (ZVT) (Oswald, 1987; Vernon, 1993). It bears some similarities to the trail making test in the Halstead-Reitan neuropsychological test battery (Reitan, 1985), however these stimuli are regularly spaced, and there is a hidden alphanumeric logical sequence to the stimuli. This task displays a 16-position matrix (4 x 4) of 1 cm. dots on the screen, spaced 3 cm. apart. The participant is required to connect the dots by clicking the mouse button on

the next correct dot in the sequence. Sequences consist of numeric or letter stimuli, increasing or decreasing, sometimes alternating. For example, a sequence could consist of numbers decreasing by one, alternating with letters increasing by two:

8, A, 7, C, 6, E, 5, G, 4, I, 3, K, 2, M, 1, O

The goal is to guess the next number/letter in the sequence and complete the trail, making as few errors as possible. When the correct dot, corresponding to the logical sequence rule is clicked, an arrow head joins the previous dot creating a trail. The time to complete each of the 10 trials, and the number of errors are recorded for each participant. This test is useful as a fluid intelligence marker (gf) and has the advantage that the participant can complete each trail by trial and error clicking, or, by intuiting the letter/number sequence thereby making fewer errors. Thus, everyone can complete the test, but those who see the patterns make fewer errors. This method removes the pressure often felt by participants when performing cognitive tests, and is similar to unobtrusive test techniques described elsewhere (e.g., Legree, Martin & Psotka, 2000). The score for this test is the number of errors made to complete each trail, thus lower error scores are related to higher fluid intelligence (gf).

Brief Evaluation of Behaviour Interview (BEB). This measure was derived from a behavioural observation of the participant's personality style demonstrated in a brief interview. The design was modelled on Ford and Tisak (1983), with the experimenter being the only judge. Open questions were directed to each participant in private, and judgements were made on a 1 to 10 scale regarding ability to speak effectively, appropriateness of nonverbal behaviours, moderate levels of eye contact, and relaxed but attentive body posture. Each participant's score on a scale of 1-10 for each behaviour was bounded by "poor or unresponsive" at the low end and "very competent or appropriate" at the high end. The scoring was based on "appropriate" behaviour for the interview, for example, a high score would be given for smiling at the appropriate times, or demonstrating appropriate affect, without being excessively emotional or devoid of emotion. Eye contact and posture were also rated as appropriate or otherwise to an interview, while the amount of fidgeting was reverse scored as it represented inappropriate behaviour for an interview situation.

For this overall score, the judge assumed that each participant had applied for a job in a customer-facing role such as retail sales, and that maturity and social skills were required for the job. The judge made the overall assessment of the participant's maturity and social competence at the end of the interview, after considering the participant's ratings on the four behaviours listed above. This score represents a behaviourally-based EI score observed in the brief interview; referred throughout this document as the brief evaluation of behaviour (BEB). BEB corresponds to the "Interview" score in the Ford and Tisak study (1983).

During the interview, a variety of questions were directed to the participant. These questions were designed to evoke an emotional response from the participant. These questions are listed below:

1. What do you want to be doing, or where do you want to be in 5-years time?
2. Do you see yourself as a leader, team player, nurturer or independent?
3. How do you tune in to the needs and feelings of your close friends?
4. How do you express warmth and concern for your friends?
5. What areas have you been a leader in?
6. How do you get people to do what you want them to do?
7. What is your birth order?
8. What do you think the role of the (substitute birth order from Q.7.) child is?
9. What was your most salient personality attribute at high school?
10. Where would you have come in a popularity poll of the class?
11. Do you believe that people have a specific purpose in life?
12. What would you like to be remembered for?

Ethics

All procedures adopted in the study were approved by the Human Ethics Committee of Macquarie University, NSW, Australia.

Apparatus

All tasks for this study were computer-based using Dell, IBM-compatible computers with 1GHz clock speed, running Windows™ 98. All computer-based tasks were developed by the author to run on IBM-compatible personal computers using Macromedia Authorware™ 7 software.

Test Procedure

Participants took the test in a quiet, computer laboratory environment. The battery of tests was planned to take 50-minutes to allow for instructions and the brief EI interview within the one-hour period. Most participants completed within the hour. Participants were given an Ethics Consent form and verbal instructions regarding the tests. This gave participants an insight to the type of tests that would follow and enabled them to withdraw at that time if they wished. The instructor demonstrated the first item of each test in the sequence to ensure familiarity and understanding. The experimenter remained with the participants for the entire test to ensure quality data, and to field queries as they arose. At the completion of the battery, each participant was interviewed briefly and thanked for their participation.

Results

The presentation of results has in some cases, such as with the correlation matrix, been limited to the tests that were found to significantly correlate with emotional intelligence (EI) in its various manifestations.

Descriptive Statistics

The following table lists the tests used in this battery, together with normative data and alpha reliabilities where available. Some tests, such as the tacit knowledge and the facial affect recognition task, were specifically developed for this study and thus there were no standardised values available to include in Table 2.

Table 2

Descriptive statistics for the battery of tests (N=103), with coefficient alpha, and standardised values from the original test-developers where available.

	This Study			Standardised Values		
Test	Mean	SD	Alpha	Mean	SD	Alpha
1. Age of participants (years)	22.25	4.17	-	-	-	-
2. BFF personality						
• Openness	44.08	8.72	.83	44.05	8.01	.80
• Conscientiousness	47.49	10.23	.91	47.47	8.56	.87
• Extraversion	48.51	8.78	.89	47.61	9.35	.90
• Agreeableness	57.36	6.17	.83	56.56	7.01	.88
• Neuroticism	36.61	8.21	.83	35.48	8.51	.89
3. Word reasoning task	.64	.15	.75	.66	.16	.70
4. Matrices task	.61	.22	.73	.44	.19	.73
5. Trail-making task	.11	.06	.81	-	-	-
6. Tacit knowledge	3.28	3.32	-	-	-	-
7. Facial affect recognition task	.55	.11	.71	-	-	-
8. Brief affect recognition task	.60	.13	.80	.65	.16	.87
9. Vocal emotion recognition	.71	.15	.63	.68	-	.62
10. Situational judgement task	.88	.41	.85	-	-	-
11. Assessing Emotions Scale	127.59	10.12	.83	128.86	15.57	.90
12. Trait meta-mood scale	109.04	11.68	.88	-	-	.86
• Attention to emotions	49.26	6.65	.82	-	-	.86
• Clarity	37.95	6.10	.67	-	-	.88
• Repair	21.83	3.54	.83	-	-	.82
13. Emotional Stroop task	-.03	.12	-	.26	-	-
14. Brief EI interview (BEB)	.61	.12	-	-	-	-

All skew and kurtosis statistics were within the range of plus one to minus one; the maximum statistic being 0.75, indicating that the data did not require transformation before further analysis (Dekker, 2005).

Test battery

The results from the tests performed by the participants produced means, standard deviations, and internal consistency reliabilities (coefficient alpha) that were very close to the standardized data supplied with the tests. This provides evidence that the tests had been administered and scored correctly, and that the Macquarie University participants were from the same general population as previous participants who had taken these tests. There were some new tests for which there was no standardized data and these will be discussed as follows.

The Tacit knowledge serial reaction time task (Nissen & Bullemer, 1987) was developed for this battery and no standardization data were available. The test however is self-calibrating as it simply measures the participant's average response time improvement on a sequence of repeated stimuli. The test is sensitive to outliers such as where a participant was distracted and did not respond for a second or more. Response times extending to more than three times the mean response time (i.e., 3 x 330 milliseconds) for all participants were rare, and these were treated as outliers and removed as they would distort the data.

The facial recognition task was developed from a subset of faces from Mazurski and Bond (1993). Although the reported recognition rates for the emotions was generally high, ranging from happiness 97.3% to fear 76%, the standardization data were not presented in terms of means, standard deviations, and alpha reliability, and so were not suitable for inclusion in Table 2.

The Situational Judgement Task (SJT) was scored on the basis of positive points for the correct emotion and negative points for incorrect emotions chosen. The correct answer, chosen in advance by the experimenter, turned out to match the modal choice by the participants, that is, a consensus was reached. This supports Legree's (1995) hypothesis that for social or emotional intelligence tests, the choice of a large sample of participants will eventually converge on the correct response (see also, Mayer, Caruso & Salovey, 2000).

Group differences

To check whether any of the tests discriminated against participants who did not have English as their first language, an independent samples t-test was run against all tests in the battery, with EFL as the grouping variable. Significant differences between participants who had English as their first language (n = 72) and participants with English as their second language (n = 31) were observed in the variables shown in Table 3.

Table 3

Tests where participants from a non-English speaking background received lower scores.

Variables	Mean score English as first language	English as second language	*t*	*p* - value
BEB	6.27	5.63	2.64	0.01
Word reasoning	14.44	13.07	1.99	0.03
Attention to emotion	50.93	45.39	4.19	0.001
TMMS global score	111.01	104.45	2.69	0.008
Voices	0.76	0.65	4.16	0.001
Extraversion	49.96	45.12	2.63	0.01

While it seems logical that English as second language would put a participant at a disadvantage in the Word analogies test, the Voices test was in a synthetic language, spoken by German radio actors, so at first glance it would appear to be culture free. Perhaps participants with English as a second language find it difficult to retrieve the appropriate English word for an emotion, thereby performing below average in Vocal affect recognition.

Gender differences

An independent samples t-test was conducted on all variables to compare scores of males and females. Significant differences between scores of females (n = 78) and males (n = 25) were only apparent in the brief evaluation of behaviour (BEB) criterion, and the Trait Meta-Mood Scale (TMMS) "Attention to emotion" factor.

Table 4

Significant differences between males and females.

Variables	Mean score Female	Male	*t*	*p* - value
BEB	6.21	5.64	2.20	0.03
Attention to emotion	50.28	46.08	2.85	0.005

Outliers

The exclusion of data as outliers in an objective manner is necessary to maintain the integrity of the data. Boxplots flag data beyond the inner fences as outliers, as these are data points

that exceed 2.575 standard deviations from the mean. These points represent the outlying 1% extreme data in the cases, 0.5 % in each tail (Howell, 1997). Heavily skewed variables with a skew statistic greater than plus or minus 1.0 can present as overly significant predictors in regression analysis and therefore must be normalised. This study adopted the Boxplot criteria and removed data if they exceeded 2.575 times the standard deviation from the mean. Variables such a tracing, stroop, and voices score had eight, two, and three outliers respectively that were removed.

Correlation matrix

As a first step in establishing covariant relationships among tests, a correlation matrix was produced using SPSS. The correlation matrix in Table 5 below includes the 14 tests of interest for this study. Table 5 depicts the correlations among all variables used in the study, with the first five variables representing the BFF personality dimensions. The next three variables, Word reasoning, Matrices and Tracing errors represent a measure of general intelligence or IQ. Tracing errors are the number of errors made in a cognitive task; thus lower scores on Tracing represent higher general intelligence. Variables from Tacit learning (9) through to the brief evaluation of behaviour (BEB) criterion score (20) are recognised tests of EI such as the AES and TMMS, or potential tests of EI such as Tacit learning, Emotional Stroop, Faces and Voices tasks. The three subfactors of the TRAIT Meta-Mood Scale (TMMS): Attention to Feelings (16), Clarity of Feelings (17), and Mood Repair (18), have been labelled Attention, Clarity and Repair for simplicity in this study.

Table 5

Correlation matrix showing the significant relationships among tests.

Variables		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1.	Openness	1																		
2.	Conscientiousness	.22*	1																	
3.	Extraversion	.13	.30**	1																
4.	Agreeableness	.41**	.45**	.38**	1															
5.	Neuroticism	.05	-.08	-.21*	-.13	1														
6.	Word reasoning	.10	-.03	.05	-.15	-.05	1													
7.	Matrices	-.08	.00	-.02	-.16	-.06	.42**	1												
8.	Tracing errors	-.08	-.15	.01	-.04	.13	-.41**	-.40**	1											
9.	Tacit learning	.16	.24*	.03	.16	.07	.11	.11	-.10	1										
10.	Faces recognition	.00	.00	.05	.01	-.12	.22*	.34**	-.24*	.13	1									
11.	Brief affect task	-.15	-.11	-.08	-.13	-.18	.13	.20*	-.14	-.02	.26**	1								
12.	Voices recognition	.17	-.01	.15	.01	-.17	.25*	.04	-.08	.19	.14	.12	1							
13.	Situational judgment	-.06	.06	-.05	.00	-.23*	.00	.27**	-.12	.12	.26**	.00	.09	1						
14.	AES global score	.27**	.31**	.38**	.57**	-.17	-.02	-.16	.00	.00	.00	.02	.17	-.09	1					
15.	TMMS global score	.09	.20*	.28**	.23*	-.28**	.15	.06	-.17	.09	.12	.08	.20*	.15	.56**	1				
16.	Attention	.05	.12	.28**	.14	.00	.11	.10	-.12	.11	.14	.06	.23*	.27**	.40**	.77**	1			
17.	Clarity	.07	.14	.13	.15	-.37**	.21*	.04	-.21*	.08	.03	.02	.08	-.01	.40**	.77**	.28**	1		
18.	Repair	.10	.21*	.20*	.25*	-.28**	-.07	-.08	.04	-.06	.10	.11	.09	-.01	.39**	.54**	.17	.27**	1	
19.	Emotional Stroop	-.07	.16	.03	.08	-.10	-.08	-.05	-.03	.07	-.03	.21*	-.12	.02	.04	.06	-.05	.10	.13	1
20.	BEB criterion score	.00	.21*	.18	.00	.00	.21*	.16	-.06	.09	.15	.10	.31**	.04	.07	.27**	.22*	.20*	.13	.03

Note: ** indicates that correlation is significant at the 0.01 level (2-tailed), * indicates significance at the 0.05 level (2-tailed).

Using the Brief Evaluation of Behaviour (BEB) as the criterion and behavioural benchmark for EI, the task that exhibits the highest correlates with BEB is the Voices recognition task, $r(101) = .31$, $p < .01$. The next most significant correlate of BEB is the TMMS global score, $r(101) = .27$, $p < .01$, followed equally by Conscientiousness and the Word reasoning task, $r(101) = .21$, $p < .01$. The Tacit learning, Situational judgement, and the emotional Stroop tasks demonstrated very low correlations with the BEB criterion score (0.09, 0.04 and 0.03 respectively) so were not considered further as potential EI predictor variables.

Having examined the zero-order correlation matrix relationships, related factors were combined into families and regression analysis was conducted to determine the ability of each test family to predict the BEB criterion.

H1: EI tests would be significant predictors of the BEB criterion

Regression analysis

To examine the predictive powers of each family of tests on the EI behavioural criterion (BEB), a regression analysis was conducted with the variance explained and the model significance tabulated below in Table 6.

Table 6

Regression analysis of the variance explained by theoretical families of tests, with the brief evaluation of behaviour (BEB) as the criterion.

Model	Predictors for BEB	β	*p* -value	Overall Model Fit	Total variance explained R^2
1. (BFF)	O	-.01	.94	*F*(5, 88) = 1.55 (*p* = .18)	8.1%
	C	.23	.05		
	E	.18	.12		
	A	-.17	.20		
	N	.03	.74		
2. (IQ)	Word	.19	.12	*F*(3, 90) = 1.67 (*p* = .18)	5.3%
	Matrices	.11	.36		
	Tracing	.06	.60		
3. ($SREI_1$)	AES	-.12	.32	*F*(2, 91) = 4.11 (*p* = .02)	8.3%
	TMMS	.34	.01		
4. ($SREI_2$)	AES	-.12	.33	*F*(4, 89) = 2.02 (*p* = .10)	8.3%
	Attention	.21	.07		
	Clarity	.17	.14		
	Repair	.09	.41		
5. (PEI)	Brief affect	.04	.73	*F*(3, 90) = 3.60 (*p* = .02)	10.7%
	Faces	.10	.36		
	Voices	.29	.01		
6.	TMMS	.22	.03	*F*(2, 91) = 7.43 (*p* = .001)	14.0%
	Voices	.27	.01		
7.	Voices	.31	.003	*F*(1, 92) = 9.66 (*p* = .003)	9.5%

Note: SREI represents self-report EI mix 1 and 2; PEI is performance EI.

Table 6, model 1 indicates that while the BFF personality factors explain 8.1% of the variance in the BEB criterion, the regression model is not statistically significant ($p = 0.18$). The

general intelligence tests in model 2 only explain 5.3% of the variance; this model also does not reach significance. This is of no concern because the personality and intelligence variables were included only to set a baseline of variance explained, and to determine if EI explains significant incremental variance.

Model 3, self-report EI model 1 ($SREI_1$), uses the AES score and the global TMMS score and explains 8.3% of the variance with a statistically significant regression model ($p = .02$). Model 4, self-report EI model 2 ($SREI_2$), includes the AES score and the TMMS subfactors to see whether the subfactors explain more variance than using the global TMMS score. No increase in variance is achieved by using the subfactors, and the regression model does not reach significance ($p = .10$).

The performance EI tests (PEI) in model 5 consisting of the brief affect recognition task (BART), the facial emotion recognition task (Faces) and the vocal affect recognition task (Voices) and explains 10.7% of the variance, with a significant regression model ($p = .02$).

The combination of the TMMS self-report EI scale and the vocal affect recognition performance EI task in model 6 explains more variance (14%) than any other combination of predictors ($p = .001$). Finally, the vocal affect recognition task in model 7, explaining 9.5% of the variance, appears to be the single best predictor of the BEB criterion of any variable in the battery ($p = .003$).

Only the EI models that reach statistical significance, 3, 5, 6 and 7, are important for this study. The combined self-report and performance EI model (6) clearly explains more variance (14%) than any of the other families of tests, while the performance EI model (4) PEI, is also a significant model and comes second explaining a satisfactory 10.7 % of the variance. These four models provide some necessary, but not sufficient support for hypothesis one; that EI tests would be significant predictors of the BEB criterion. To fully support hypothesis one, it must be shown that EI can explain significant incremental variance above and beyond the traditional constructs of personality (BFF) and general intelligence (IQ). To complete hypothesis one testing, hierarchical regression analysis was used.

Hierarchical regression analysis

Hierarchical analysis was conducted to determine the ability of EI tests to explain additional and statistically significant variance in the BEB criterion after controlling for personality factors and general intelligence.

Step 1 in Table 7 is the first model in the hierarchical regression and indicates that a variance of 8.1% in the BEB criterion can be attributed to personality factors. Step 2 adds the general intelligence (IQ) scores to the personality predictors into the regression equation increasing the variance explained by 4.7% up to 12.7%. Neither models 1 or 2 reach statistical significance ($p < .05$) with respect to the incremental variance explained.

Model 3 adds the predictor variables from the two self-report EI tests, AES and TMMS, plus the performance EI tests: the brief affect recognition task, the facial recognition task, and the vocal affect recognition task. The final model (3) consisting of five EI tasks explains a statistically significant 13% of variance above and beyond personality and IQ tests (p = .02). This final analysis, explaining an incremental 13% of variance, provides the final piece of evidence required to accept hypothesis one; that EI tests can indeed explain a significant amount of variance in the BEB criterion, above and beyond the traditional constructs of personality and IQ.

Table 7

Hierarchical regression analysis of families of predictors entered at each step, for the criterion, the behavioural measure BEB.

Step	Predictors for BEB	β	*p* -value	Overall Model Fit	Total variance explained R^2	Incremental variance above step 1 ΔR^2	Incremental variance significance *p* - value
1.	O	-.01	.94	*F*(5, 88) = 1.55 (*p* = .18)	8.1%	-	-
	C	.23	.05				
	E	.18	.12				
	A	-.17	.20				
	N	.03	.74				
2.	O	-.04	.73	*F*(5, 85) = 1.55 (*p* = .15)	12.7%	4.7%	0.22
	C	.24	.05				
	E	.15	.19				
	A	-.10	.47				
	N	.04	.68				
	Word	.19	.12				
	Matrices	.10	.38				
	Tracing	.08	.50				
3.	O	-.07	.52	*F*(13, 80) = 2.13 (*p* = .02)	25.7%	13.0%	0.02
	C	.25	.03				
	E	.10	.36				
	A	-.04	.79				
	N	.15	.17				
	Word	.10	.42				
	Matrices	.09	.46				
	Tracing	.11	.33				
	AES	-.16	.26				
	TMMS	.27	.03				
	Brief affect	.06	.55				
	Faces	.05	.65				
	Voices	.27	.01				

H2: The magnitude of the variance explained would justify further research

It was noted in the introduction that at this advanced stage of psychometric research, any new construct that explains more than 5% of additional variance over existing constructs is worthy of further research (Mayer et al., 2000b). This study has demonstrated an additional 13% of variance that can be explained by a combination of self-report and performance EI tests. This magnitude exceeds the benchmark set by the major researchers in the discipline of EI.

Thus hypothesis two is unequivocally confirmed; that the magnitude of variance explained by EI tests justifies further research into the construct.

H3: The tests would cluster into theoretically defined groups

To isolate related constructs within the test battery, an exploratory factor analysis was conducted using SPSS. The results of the analysis are shown below in Table 8.

Table 8

Principal Components analysis with orthogonal (Varimax) rotation of test results indicating clustering of related constructs. For clarity, loadings below .30 are omitted, and salient loadings above .30 are shown in bold.

Variables	Component			
	1	2	3	4
Openness	**.66**			
Conscientiousness	**.67**			
Extraversion	**.38**			
Agreeableness	**.76**			
Tacit learning score	**.41**			
Word reasoning		**.69**		
Tracing errors		**-.75**		
Matrices		**.75**		
Facial affect recognition		**.52**		
Brief affect recognition	-.39	**.32**		
AES	.47		**.59**	
Repair			**.66**	
Clarity			**.59**	
Neuroticism			**-.71**	
BEB				**.69**
Attention				**.59**
Vocal affect recognition				**.70**

The test results clustered into a BFF personality component (1); a general intelligence component (2); a self-report EI component (3); and a performance EI component (4). The only loading that was unexpected was the brief affect recognition task loading negatively (-.39) on the

BFF personality component. The AES loading on the BFF personality component (.47) is not unexpected as self-report EI tests have typically been found to be closely related to personality (e.g., Davies et al., 1998; Petrides et al., 2007; Bracket & Mayer, 2003, Brackett, Rivers, Shiffman, Lerner & Salovey, 2006; Matthews et al., 2002, 2004). These anomalies were examined for significance using Confirmatory Factor Analysis (CFA), reported in the following section.

This Principal Components model explained 49.2% of the variance of the tests, indicating reasonably high interrelatedness among the tests (Child, 1976, 2006). Communalities, that is, shared variance among the tests, were all greater than 0.3, ranging from 0.33 to 0.70, confirming that in general, that the tests were reliable (Child, 2006). The scree plot shown in Figure 1 indicated the possibility of a 6-factor model, although the eigenvalues of the fifth and sixth factor were borderline (1.06 and 1.01), and were therefore dropped. Thus, a 4-factor model was chosen as this was congruent with the four theoretical families of tests being administered.

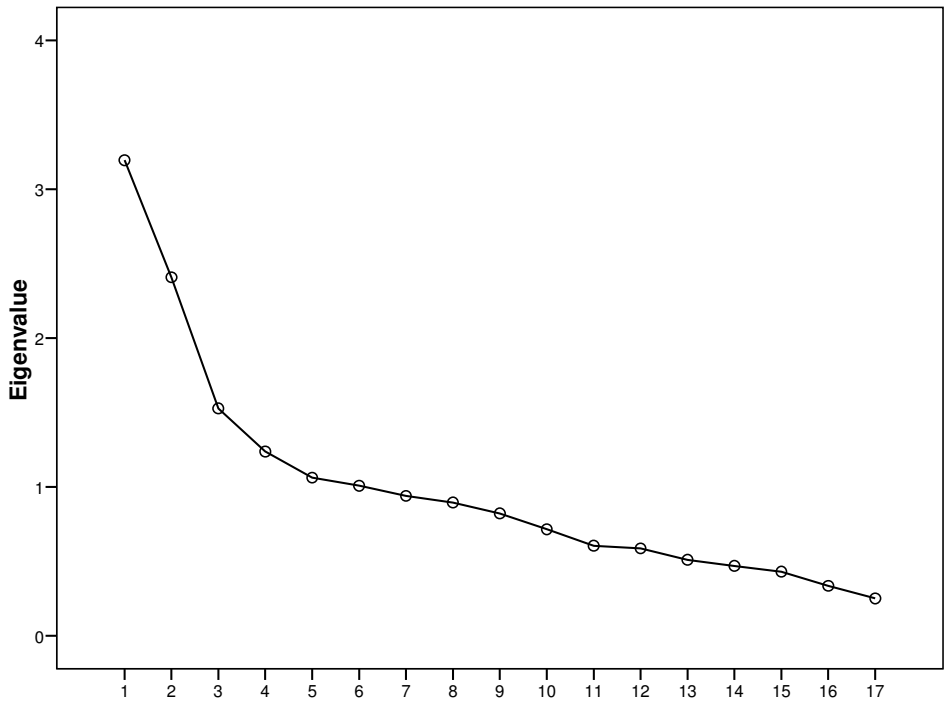


Figure 1

Scree plot showing why a 4-factor solution was chosen; factors 5 and 6 are close to unity.

Where there were loadings greater than 0.3 on factors that did not appear to be theoretically related, these loadings will be validated or removed in the confirmatory factor analysis (CFA) model.

Factor one in Table 8, explaining 13.3 % of the variance, represents four of the BFF personality constructs (Openness, Conscientiousness, Extraversion and Agreeableness) plus the

Tacit learning score. In the zero-order correlation matrix (Table 5), Tacit learning had only one significant correlation, and that was with Conscientiousness; thus the Tacit learning test appears to be a proxy for Conscientiousness. Tacit learning loads on the same factor as Conscientiousness, and has moderate loadings on all factors, indicating that it does not cluster clearly with any of the theoretical group of tests. Tacit learning was therefore dropped from further analysis because its variance was spread across all factors, and it did not appear to belong with the performance EI group of tests as expected (factor 4).

Factor two in Table 8, explaining 13.1% of the variance, represents academic intelligence or IQ, with a bias toward fluid intelligence as indicated by the highest loadings of 0.75 and -0.75 on the Matrix and Tracing tasks. Note that the Tracing task has a negative correlation because it is a measure of the tracing errors made by the participant.

Factor three in Table 8, explaining 12.7 % of the variance, has Neuroticism as the highest negative loading (-.71), with TMMS Repair the next highest loading of 0.66. Neuroticism has been found to be negatively related to self-report EI scores in other studies (e.g., Schutte et al., 1998; Bar-On, 2000; Brackett & Mayer, 2003), thus providing some construct validation for these self-report EI tests. The AES and the TMMS subfactor Clarity, also have high loadings on this factor confirming that it is a self-report EI factor.

Factor four in Table 8, accounting for 10.2 % of the variance demonstrates that the Vocal affect recognition task has the highest loading (0.70) on this factor. The brief evaluation of behaviour (BEB) also falls under this factor with a substantial loading of 0.69. The ability to discriminate among vocal emotions also has a significant positive correlation (0.31, $p < .01$) with the BEB criterion in the zero-order correlation matrix (Table 5), thus confirming a close relationship between these variables. The TMMS Attention to feelings score also loads on this factor, thus completing a cluster of measures that, apart from the Attention self-report task, tend to indicate a performance EI factor.

An alternative PC component analysis using oblique rotation (Oblimin) yielded the same factor structure with only minor factor loading differences. This was interpreted as evidence that the structure of the factors shown above was robust and stable. The orthogonal rotation results shown above in Table 8 therefore appear to be a reasonable representation of the factorial structure of the data, sufficient to proceed to a confirmatory factor analysis (CFA).

Confirmatory factor analysis

Exploratory factor analysis (EFA) provides insufficient evidence for claiming underlying structure and postulating the existence of latent variables such as general intelligence, EI, and personality facets. Confirmatory factor analysis (CFA) is one of the functions available from structural equation modelling (SEM) computer packages such as AMOS (Arbuckle, 2005, 2005a)

that address this uncertainty. SEM extends the general linear model (GLM) by enabling the simultaneous testing of multiple regression equations (Arbuckle, 2005). CFA can be used to perform hypothesis testing on models such as those produced by EFA, to test if the data fits the model. If there is a significant model fit with the data, as determined by the fit statistics cut-off values shown below, this constitutes evidence to support the model. The use of CFA is an example of applying falsification theory (Popper, 1963) to critically and scientifically evaluate models and theoretical structures.

The following CFA shown in Figure 2 was produced using AMOS. The cut-off values used to determine SEM model fit are taken from the AMOS 6.0 User's Guide (Arbuckle, 2005) and consist of:

1. The *chi-square* test of model fit should be greater than 0.05.
2. A *chi-square/df* (CMIN/DF) ratio of less than 3 indicates an acceptable fit.
3. A Tucker-Lewis Index (TLI) close to 1 indicates a very good fit.
4. RMSEA of less than 0.05 indicates a close fit.

All fit indices for the model indicate a good fit except that the TLI may be a bit low considering that researchers typically use 0.95 as the cut-off, however 0.90 is considered adequate (e.g., Bentler & Bonett, 1980, cited in Onwuegbuzie & Daniel, 2003) and the AMOS 6.0 User's Guide only proffers "close to 1" as guidance (p. 503). The TLI fit index aside, three out of four fit indices indicate a good fit, and the RMSEA indicates a close fit, thus the model is deemed to be an acceptable representation of the underlying data. The overall data model indicates a four-factor model of abilities clustering into the expected theoretical domains as follows: personality, self-report EI, performance EI, and general intelligence.

The correlation of 0.69 between the BFF personality factor and the self-report EI factor indicates a strong association, while the correlation of 0.48 between performance EI factor and the general intelligence factor represents another reasonably strong relationship, thus providing confirmatory evidence for trait and performance EI (e.g., Mayer et al., 2000; Matthews et al., 2002; Petrides & Furnham, 2001). These two factors are divergent as represented by a low correlation of 0.26 (c.f. similar results in Brackett & Mayer, 2003).

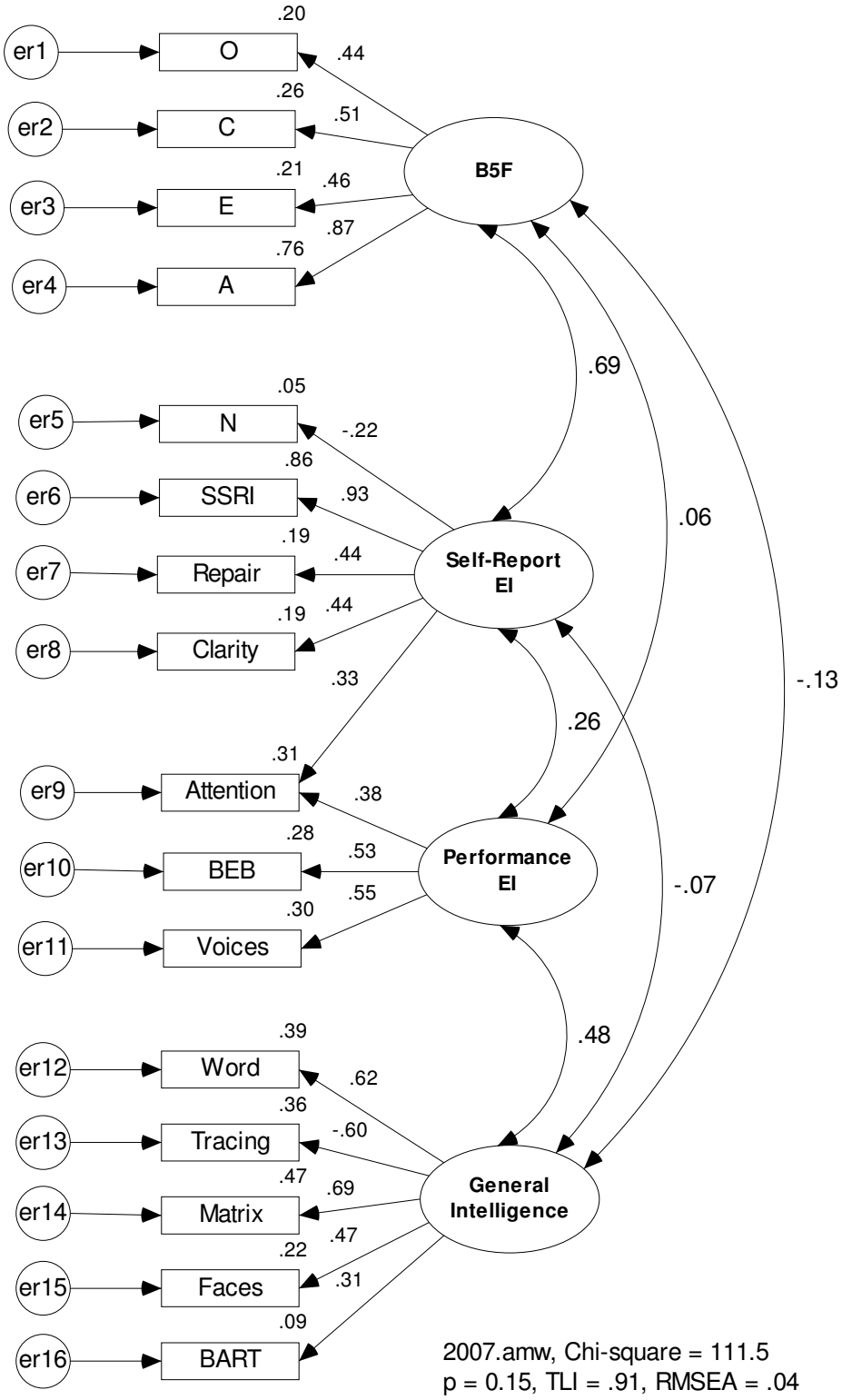


Figure 2

Confirmatory factor analysis supporting a 4-factor solution for the data consisting of: BFF personality, self-report EI, performance EI, and general intelligence latent factors. The statistics indicate a good fit for the model (p = .15, *CMIN/df* = 1.15, *TLI* = .91, *RMSEA* = .04).

Discussion

The three main aims of this study were to determine: (i) whether EI test scores would be significant predictors of a behavioural criterion, (ii) whether the magnitude of EI variance explained (controlling for personality and IQ) would provide sufficient justification for further research, and, (iii) to examine whether EI complies with the structures that have been proposed by EI theorists. The following discussion will focus on the implications and findings in the light of these aims.

The first hypothesis proposed that EI test results would be significant predictors for the behavioural observation variable, BEB. This hypothesis was based on the proposition that individuals who exhibit *maturity* and good *character* are high in EI (Goleman, 1995, 1996), and are therefore more successful in life and business. Within the limitations of the EI tests and the behavioural criterion utilised, hypothesis one was clearly confirmed, with EI tests explaining more variance than any other family of tests.

Incremental variance explained by EI

Turning to the hierarchical regression analysis, the EI factors: AES, TMMS, BART, Faces and Voices explained an incremental variance of 13% over the personality and general intelligence predictors. This provides the final piece of evidence required to support hypothesis one; that EI tests would be significant predictors of the behavioural criterion (BEB), when controlling for personality and general intelligence. The only statistically significant EI predictors in step 3, Table 7, are the TMMS and the Voices task, which enable the model to be simplified further if required.

However, the important question is whether this incremental variance of 13% attributable to EI testing is significant in a real-world sense; sufficient justification for more EI research. Research has established that IQ accounts for 20% of the factors that determine life success (Goleman, 1995), while others state a broader range of 10% to 30% of variance for IQ predicting job performance (e.g., Matthews et al., 2002). Of the remaining 70% or more variance potentially available for other factors including EI to explain, it seems that the 13% explained by EI falls well short of the 70% of the available variance. It is understandable in this study that IQ only explained 5% of the criterion because it was not job or life success related. However, the criterion was based on *character* and *maturity*, so EI tests would be expected to explain more variance (c.f., Goleman, 1995, 1996). Overall, it is not surprising that the broad range of skills that comprise *character* and *maturity* could ever be measured with 63-items of self-report tests that comprise the AES and TMMS. In this regard, 13% of the variance is a reasonable starting point for EI research, and encouragement to develop additional tests that will explain additional variance. Perhaps the fleshing out of the additional 70% of the variance will require tests that include, “genuine situations with real people” as proposed by Thorndike last century (Thorndike, 1920, p. 231).

On the other hand it is important to consider how 13% of variance explained compares with the findings of other experienced researchers in the EI field. In this regard Bradberry and Greaves (2004) reported that the MSCEIT explained 6% of the variance in job performance for 36 senior leaders. As mentioned previously in the results, researchers in the EI field consider that 5% of additional variance over existing constructs is worthy of further research (Mayer, Salovey & Caruso, 2000b). In this regard this study's achievement of explaining 13% of the variance, 2.5 times the 5% benchmark for variance set by Mayer et al., represents a worthy contribution to EI research. With enhancements and extensions to the methodology utilised here, the ability to explain even greater amounts of variance is clearly possible.

Construct validation

The principal components factor analysis (PCA) and the confirmatory factor analysis (CFA) demonstrated expected clustering of tests, but also some unexpected results. There was clearly a factorial divide between the self-report or trait EI tests and the performance EI tests as demonstrated in the PCA and the CFA results. This adds confirmatory evidence to the trait versus performance EI theory divide (Mayer et al., 2000; Petrides & Furnham, 2001; Matthews et al., 2002). However, the Faces and brief affect (BART) task aligned with the general intelligence factors, and had minimal loadings on the performance EI task in the PCA and CFA analyses. This confirms other research findings where facial recognition tasks aligned with general intelligence, rather than the EI, factor (Roberts, Schulze, O'Brien, MacCann, Reid & Maul, 2006). An alternative theoretical approach would be to have the facial emotion recognition ability removed from EI constructs, and classified under Carroll's (1993) taxonomy of cognition, stratum II constructs: General memory and learning (2Y), or broad visual perception (2V). These subsume stratum I abilities of Visual memory (MV) and Imagery (IM) respectively.

The implications of this study for performance EI tests such as the MSCEIT are profound. The MSCEIT representation of EI has been purported to be closely related to IQ, based on higher correlations than trait EI tests with general intelligence markers (Matthews et al., 2002; Brackett & Mayer, 2003). In this study, situational judgement tasks (SJT), similar to those comprising 25% of the MSCEIT, correlated significantly with the Matrices task ($p < .001$). The facial emotion recognition tasks in this study, similar to those comprising 12% of the MSCEIT, were found by factor analysis to be classified under the general intelligence factor. Thus, this study suggests that one third (37%) of the MSCEIT items could be considered proxy measures of general intelligence, rather than EI as proposed by the authors (Mayer et al., 2002; Reid, 2002). It is therefore no surprise that the claim that performance EI appears to be related to general intelligence is often made (Mayer et al., 1999; Matthews et al., 2002). It is related because it appears to contain tests that are proxies

for general intelligence. In this regard it is not extending EI research, but simply expanding the general intelligence constructs.

By contrast, the vocal affect recognition task (Voices) correlated significantly with the BEB criterion ($p < .01$) thus forming a performance EI (PEI) factor. Voices had a significant correlation with the word reasoning task representing crystallised intelligence (gc) ($p < .05$), but was uncorrelated with Matrices representing fluid intelligence (gf). Since the Voices task had the highest correlation with the BEB criterion, a prime position that it maintained even when controlling for other variables, it must be considered the most predictive task for this application, and worthy of further investigation as part of the EI realm.

Evaluating the tests

The global TMMS score predicted the same variance as the separate subscales, Attention, Clarity and Repair, indicating a well-designed structure that could be suitably represented by a unitary EI factor. The AES had a low but positive zero-order correlation (.07) with the BEB criterion, but this turned to a negative beta coefficient in the regression analysis. The beta weightings suggests that the AES played a secondary role to the TMMS and became a suppressor variable in the final regression. While a suppressor variable reduces error variance in the regression, it is unfortunate that one of the main EI tests performed in this capacity, resulting in a negative relationship with the criterion when controlling for all other variables. The conclusion is that the TMMS is the more effective of the two EI scales to use for this application although the reasons why are less obvious.

One difference between the TMMS and the AES is that the TMMS is effectively counterbalanced with 16 out of the 30 items reverse scored, while the AES only has 3 out of the 33 items reverse keyed. The preponderance of the positive response directionality in the AES could potentially create an acquiescent response style that is typically considered to be problematic in tests such as the 16-PF (Cattell, 1994). The intrusion of common variance due to acquiescence would potentially detract from the specific variance extracted by each item. In addition, the TMMS has a 3-factor structure (Attention, Repair and Clarity), while the 33-items selected out of the original pool of 62-items for the AES, were chosen on the basis of a unifactorial loading. However, in subsequent studies with similar sample size to that used by Schutte et al. (1998), the AES exhibited a 4-factor structure (Saklofske, Austin & Minski, 2003; Petrides & Furnham, 2000a). Thus, the main difference between the AES and the TMMS appears to be the preponderance of positively-keyed items in the AES. A detailed item analysis and further research would be required to determine why the TMMS demonstrated better predictive validity than the AES.

It must be noted from the correlation matrix that some tests did not explain enough variance in the BEB criterion to warrant further consideration. The Tacit learning scale, a potential test of

practical intelligence (e.g., Hedlund & Sternberg, 2000), did not have any significant predictive properties in any of the regression models examined. The emotional Stroop task had no impact in any of the regression models. Stroop results were problematic and contained many outliers, where participants appeared to have a break from the test for seconds at a time. When trying to measure response time differences in the tens of milliseconds, this erratic behaviour rendered the results virtually useless. The situational judgement task (SJT) provided no significant effect in the regression analysis which was surprising considering that the Managing Emotions factor, the Mayer et al. version of situational judgement tests, comprises 25% of the MSCEIT; sections D and H. The SJT correlated significantly with Matrices and Faces indicating that it fits better into the general intelligence structure.

Limitations and future directions

After reviewing how this study could be improved, three areas will be discussed. First, the use of the experimenter as the single rater of the behavioural criterion (BEB) was problematic because of issues arising while managing the participants as they finished the computer-based tests that placed a strain on simultaneously conducting interviews. On the positive side, there is consistency in ratings style with a single rater; however, a single rater may not be aware of the social ability nuances of the various participants, and using multiple raters would provide a more complete picture. Nonetheless, multiple raters introduce the problem of interrater reliability, and therefore are not a simple solution to the rating consistency problem. Furthermore, it is difficult to rate students from non-English speaking backgrounds using the same cues that have been found to be applicable to Western students, because the expression of personality attributes, social maturity and good character, vary from culture to culture. The lower score on the BEB criterion for participants without English as their first language points to potential bias in this rating.

Second, the OCEANIC BFF personality test demonstrated excessive intercorrelations among factors, especially between Agreeableness and Openness, Conscientiousness and Extraversion. Significant correlations between these factors, ranging from 0.38 to 0.45 ($p < .01$) could explain why a latent BFF personality factor arose in the PCA and the CFA model. The use of a more established BFF personality inventory such as the NEO (Costa & McCrae, 1985), Goldberg's Big-Five markers (1992, 1994) or Saucier's (1994) mini-markers would potentially overcome this high intercorrelation and add credibility to the study.

Third, the TMMS and the AES are experimental EI test batteries, and results cannot be directly extrapolated to the most widely used and commercially available EI tests, the MSCEIT (Mayer et al., 2002c, 2002d), and the EQ-i (Bar-On, 1997, 1997a). EI is still a new and poorly defined construct; to a large extent it is defined in terms of the current MSCEIT and EQ-i

operationalisations of the construct. This raises the criticism that was originally applied to general intelligence as simply being, "what the tests test" (Boring, 1923) - a somewhat circular definition.

Conclusions

This study attempted to use EI testing to predict mature individuals of good character as may typically be sought for employment as operational staff, and potentially leadership positions. The moderate success that was obtained using EI measures, explaining 13% of the variance, exceeded intelligence and personality as predictors. This magnitude of variance explained by EI surpasses the 5% benchmark for valuable new constructs, set by leaders in the EI research area (Mayer et al., 2000b). Thus, the EI concept looks promising as a new measure; however, it still does not reach the hyperbolic heights of explaining up to 80% of the variance suggested by its main advocate, Daniel Goleman (1995, 1996, 1998, 2001a, 2001b, 2002, 2003, 2006). This research does provide encouragement for further studies using a similar criterion and a wider range of commercial EI tests. This methodology appears to be able to provide direction and potential advantages in terms of tools and guidelines for superior staff selection processes.

In the next study the problems identified in the limitations section above were addressed wherever possible. This entailed the use of a different personality inventory to reduce the high intercorrelations between the factors in the OCEANIC personality inventory, and the use of multiple raters for evaluating the behavioural interview. The employment of multiple raters would provide an opportunity for calculating the reliability of the brief evaluation of behaviour (BEB), and to provide a more objectively grounded measure. Finally, it was decided to include an academic performance criterion in the study to determine if EI scores would predict academic as well as behavioural outcomes.

Chapter 3

The results of study 1, discussed in chapter 2, were promising in that they demonstrated that a collection of emotion-related tests were able to explain a significant proportion of the variance in a behavioural criterion of emotional maturity. However, the variance explained by the EI tests fell far short of the 80% of variance that Goleman (1995) had indicated that EI could account for. The shortfall in the variance explained may have been due to some of the previous study's measures that were found to have shortcomings. With these considerations in mind, a second study with improved measures and controls was conducted using a similar methodology to study 1.

In the early 1900s, Binet noted that intelligence was of the utmost importance for success in everyday life. Binet's broad definition of intelligence encompassed judgment, reasoning, initiative, adaptation and comprehension (Binet & Simon, 1905, cited in Mackintosh, 1998). However, when the first Stanford-Binet test was published in 1911, it consisted of a narrow range of tests such as counting, digit-span, mental arithmetic, arranging and rhyming; factors that are normally associated with academic intelligence test batteries for IQ testing. Binet's broader definitions including judgment and initiative seemed to have been lost along the way, perhaps due to the difficulties of developing reliable scales for these areas, and also due to education authorities being more interested in testing formal mathematics and language abilities (Mackintosh, 1998).

By 1920 it was clear that the definition of intelligence was becoming increasingly narrow. Thorndike proposed broadening the definition to include both practical and social intelligence (SI), in addition to academic intelligence. His original definition for Social Intelligence was "… the ability to understand and manage men and women ... to act wisely in human relations" (Thorndike, 1920, p. 228). The concept of SI was intuitively appealing, however, the isolation of SI as a discrete factor was generally unsuccessful in the ensuing years (Kihlstrom & Kantor, 2000) except for a notable study by Ford and Tisak (1983).

Wechsler argued that intellectual ability, as it applied to adaptation to life's challenges, should not be confined to logical and abstract perception, but should be regarded as the expression of the entire personality combined with mental abilities. To distinguish between academic intelligence (IQ) and the personality-related factors he introduced the term "non-intellective factors of intelligence" (Wechsler, 1940; 1950, p. 78). Although Wechsler acknowledged a broader definition of intelligence, he did not subscribe to Thorndike's concept of social intelligence, and in a later paper stated that, "social intelligence is just general intelligence applied to social situations" (Wechsler, 1958, p. 75). Clearly, Wechsler believed that there were other factors affecting intelligent adaptation to life's challenges, above and beyond what IQ testing covered, but he stopped short of proposing an SI factor that was distinct from the general intelligence factor. For

example, there may be individual "social intelligence" scales that test appropriateness of behaviour in social situations (e.g., Legree, 1995); however, results from these and other similar tests may not correlate or cluster to produce a common SI factor as is required to postulate a general intelligence factor (g) (Mackintosh, 1998).

The interest in EI generated by Goleman (1995) led to the publication of a variety of commercial EI tests: the Multifactor Emotional Intelligence scale (MEIS), (Mayer, Salovey & Caruso, 1997) the Mayer, Salovey and Caruso EI test (MSCEIT), (Mayer, Salovey & Caruso, 2002c, 2002d); and the Bar-On Emotional Quotient Inventory (EQ-i) (Bar-On, 1997). These tests were developed from differing theoretical standpoints: the MEIS and the MSCEIT based on a model that assumes that there are correct answers to questions about emotional-cognitive processing (e.g., Mayer et al., 2002c, 2002d); and, the EQ-i that arose from Bar-On's PhD thesis regarding the factors that contribute to satisfaction in life (Bar-On, 1997, 1997a). The MSCEIT is defined as an ability test because it assumes there are correct answers to emotion processing and situational judgement questions, while the EQ-i is a self-report questionnaire with no right or wrong answers, comprising 15 subscales and referred to as a "mixed EI test" (Mayer, 2000). Scores on the MSCEIT and the EQ-i are correlated at 0.21 representing a shared variance of only 4% (Brackett & Mayer, 2003). This very small overlap has led theorists to propose that there are two kinds of EI: ability EI, maximum performance, or information processing EI with the MSCEIT as the prime example, and trait, self-report, or mixed EI for the EQ-i and other self-report EI measures (Petrides & Furnham, 2001, 2003, 2004).

Research into the effectiveness of ability EI as a predictor of performance has been limited due to the MSCEIT being the only major offering in this area, with complex scoring keys due to the consensus scoring algorithm utilised (e.g., Legree, 1995). Being a proprietary instrument, the scoring keys are unpublished and unavailable; thus rendering it of limited usefulness in a teaching and research environment. Furthermore, Bradberry and Greaves (2004) cite a study in which the 28-item EIA self-report, or trait, EI test explained twice as much variance in job performance of senior managers (13%) than did the 136-item MSCEIT. It is for these reasons that self-report, trait EI tests were chosen for use in this study.

EI as a unified intelligence

An additional consideration for this study was to determine whether global EI is a high-level construct representing the four subfactors (Goleman, Boyatzis & Mckee, 2002) in a similar way that *g* represents general intelligence; the next level up from the subfactors of fluid intelligence, crystallised intelligence, general memory, visual perception, and so on (see Carroll, 1993). Most EI theories assume that a global EI factor exists, and accordingly sum the subfactors to produce a total EIQ score (Mayer et al., 2002c), or a total EQ score (Bar-On, 1997).

Spearman's (1904) proposal that there is a general intelligence factor originated from the observation that individual test scores such as numeric, spatial, memory, and comprehension were positively correlated. The explanation for the common variance was that it was due to a global general intelligence factor (*g*) (Mackintosh, 1998). This concomitant variation (Mill, 1843) is often referred to as the "positive manifold," whereby all mental ability tests tend to correlate positively with one another (Jensen, 1981, cited in Macintosh, 1998). On the other hand, Thurstone believed that the positive manifold was due to imperfect tests with common method variance such as being pencil-and-paper based. Thurstone refined his IQ tests and used factor rotation techniques to identify six primary mental ability factors such as numerical, verbal, and spatial reasoning that reduced intercorrelations to negligible proportions; virtually eliminating the common factor. This raises the question of whether EI can be represented as a global EI score, similar to Spearman's *g*, or whether it is a collection of primary mental abilities as Thurstone proposed; similar to a non-intellective array of loosely-related factors (Wechsler, 1950).

From a theoretical standpoint, the practice of adding EI subscales may be questioned on the grounds that they have not been proven to be additive structures. Temperature is a good example. When two liquids are added together, the resulting temperature is not the sum of the two original temperatures (Michell, 1990). Thus, one aim of this study was to shed some light on the subject of whether global EI scores, produced by summing the subfactors, adequately represent the underlying subscales; or whether the subscales behave as a loosely-related set of abilities and non-intellective factors. Whether or not this common procedure of summing subfactors has empirical merit will be one of the issues explored in this study.

EI tests used in this study

The overall aim of this study was to determine the predictive validity of EI tests; accordingly, two self-report or trait EI tests were selected for evaluation. The first EI test used in this study is the Trait Emotional Intelligence Questionnaire Short Form (TEIQue) consisting of 30 items (Petrides, Perez, & Furnham, 2003, 2004). The other is the Emotional Competence Assessment (ECA) based on Goleman's 4-factor definition of EI that forms the basis for the EIA mentioned above (Goleman et al., 2002). The four factors are: "self-awareness, self-management, social-awareness and relationship-management" (Goleman et al., 2002, p. 253). These four factors are typically summed to produce a global EI score. The TEIQue is freely available from the authors for research, and the ECA is the product of three-years of development at Macquarie University and administration to over 1500 students (Reid, 2006). Due to the use of the common Goleman model, it was assumed that the ECA would exhibit similar predictive properties to the EIA discussed above (see Bradberry & Greaves, 2004).

Petrides et al. (2007) used the Emotion Control Questionnaire, the Satisfaction with Life Scale, the Coping Styles Questionnaire, the Revised Self-Monitoring Scale, and the Aggression Questionnaire as criteria for the TEIQue EI test. These self-report scales that are compared with the TEIQue EI test, also a self-report scale, create a comparison between self-report questionnaires that potentially has limited validity. Self-report questionnaires are prone to bias and exaggeration (Matthews et al., 2002), and as such are less satisfactory than behavioural criteria such as performance ratings by a manager, or behavioural observations by a third-party. Brackett and Mayer (2003) claim that their criterion measures: daily smoking behaviour, illegal drug use, alcohol consumption, and social deviance are more behaviourally-based as they are "observable and potentially verifiable" (p. 1151). However, self-reporting on alcohol, illegal drug usage, and social deviance still enables the participant to under-report if they feel so inclined, and thus retains the potential problem of bias inherent in self-reporting. On the other hand, academic performance as measured by grades, ranks, and GPAs are criteria that can be accessed in an academic environment and do not suffer from the problems of self-reporting (e.g., Schutte et al., 1998). Overall, apart from self-report inventories, there is a limited choice of objective criteria that can be related to EI scores when testing first-year psychology students. It was hoped that the two objective performance criteria used in this study, a brief evaluation of behaviour (BEB), and first-semester academic achievement (PSY104), would provide a useful grounding for EI theory.

In summary, this study utilised a similar methodology to Study 1 (Chapter 2), but with the objective of obtaining more valid outcomes through the use of improved EI scales, multiple judges for the behavioural criterion, an improved BFF personality scale, and additional measures such as motivation, impulse control and empathy scales. In addition, an academic performance criterion was added to determine whether EI has the potential to predict first-year university performance. This study also examined whether the new EI tests would support a unitary construct with a global EI score.

Summary of aims and hypotheses

The overall aim of this study was to replicate the Ford and Tisak (1983) findings and establish whether putative tests of EI can predict behavioural and academic outcomes. The overall aims can be summarised in three formal hypotheses:

H1: If EI is indeed a unified intelligence that can be represented by a single global factor, then correlation matrix patterns and structural equation models (SEM) will confirm this structure.

H2: If EI is a new and useful psychological construct, then it will explain significant incremental variance above and beyond existing personality measures, when predicting emotionally intelligent behaviour derived from a behavioural interview (BEB).

H3: If EI is a new and useful psychological construct, then it will explain significant incremental variance over existing intelligence and personality measures, when predicting first-semester psychology academic results (PSY104).

Method

Participants

One hundred and fifteen first-year psychology students (84 female - 73%), from the School of Psychology, Macquarie University, took part in this study as part of their first-year course requirements. Their ages ranged from 18.2 to 55.7 years, with a mean of 22.57 years, SD = 7.55. The country of birth of 78 of the participants was Australia or a Western-speaking country (68%), with 37 (32%) from non-English speaking countries, predominantly Asian, but with some born in Europe and the Middle East.

Biographical Questionnaire

The biographical questionnaire was designed to collect the following information:

1. Gender
2. Date of birth (optional)
3. First language
4. Country of birth
5. Study major

Tasks in this Battery of Tests

A wide variety of tasks designed to cover a range of psychological domains were administered to establish convergent and divergent validity with Emotional Intelligence. The experimental battery consisted of the following tasks shown in Table 9.

Table 9

Tasks performed by participants in this study to evaluate both test scores and behavioural style.

Task	Type	Author/Source
1. Biographical Details (BIO)	General	
2. Brief Affect Recognition (Faces)	Facial Affect	Ekman, 2003
3. Tracing	Gf	Oswald, 1987
4. Personality questionnaire (PACL)	BFF	Saucier, 1994
5. Word reasoning (Word)	Gc	Stankov, 1997
6. Tacit Knowledge Learning task (TL)	Tacit	Nissen, 1987
7. Motivation Questionnaire (Motivation)	Motivation	Misc. sources
8. Impulsivity Scale (Impulse Control)	Impulsivity	Tellegen, 1982
9. Empathy Scale (Empathy)	Empathy	Davis, 1980, 1984
10. Mach IV scale (Mach IV)	Mach	Christie & Geiss, 1970
12. Trait EIQ (TEIQue)	EI	Petrides, 2002
13. Emotional Competence Assessment (ECA)	EI	Reid, 2005
14. Brief Evaluation of Behaviour (BEB)	EI – Interview	Ford & Tisak, 1983

Brief affect recognition task (Faces). The Micro Emotion Training Tool (METT) was administered as a test of facial affect recognition ability (Ekman, 2003, 2003a). Facial emotion recognition ability is a component of EI (Mayer et al., 2002c). The METT utilises the same stimuli as the Japanese and Caucasian brief affect recognition task (JACBART) (Matsumoto et al., 2000). The METT comprises 56 stimuli and present a brief one-fifteenth of a second emotion, forward and backward-masked by the same face without the emotional expression. Internal consistency reliability of the test (Cronbach's alpha) is 0.86 for $1/15^{th}$ of a second exposure (Matsumoto et al., 2000). After each stimulus is presented, the participant is asked to click a button to rate the expression as one of the following seven emotions: *Happiness, Disgust, Sadness, Anger, Surprise, or Fear*. A high score in Faces indicates a superior ability to recognise emotion, and thus contributes to a higher total EI score. Contempt items were removed from the stimuli presented, as a previous study had indicated that there was general confusion between Contempt and Disgust emotions, which lowered the overall reliability of the test (Reid, 2004).

Tracing task. The tracing task is a trail-making type of test (Roberts, 1998); Der Zahlen-Verbindungs test (ZVT) in German. This test was included to provide a marker of fluid intelligence, to establish the convergent and divergent validity of EI. In this version of the task, participants are required to connect the dots in a four by four layout of number and letter stimuli, and thus complete the trail from beginning to end with as few mistakes as possible. The four by four screen layout of dots containing the letters or numbers was ten by ten centimetres. When the correct

dot (corresponding to a logical sequence rule) is clicked, an arrow head joins the dots, thus creating a trail. Tracing tasks such as the Halstead-Reitan trail-making test assess cognitive efficiency and functioning of both hemispheres, and thus are sensitive measure of cognitive ability (Vickers, 1998). The time to complete each of the ten trials, and the number of errors, are recorded for each participant. This task is used as a marker of fluid intelligence (Gf) as it was found to exhibit convergent validity, a correlation of 0.43, with a (Raven's-type) Matrix task used in a previous study by the author (Reid, 2004).

Personality Questionnaire. This 40-item scale of "Mini Markers" (Saucier, 1994) is a brief version of Goldberg's (1992) Big-Five markers. It is based on the "Big-Five" factor (BFF) model (McCrae & John, 1992) and provides a measure of the five global factors, Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Personality factors were included to determine whether EI is a divergent construct from personality (e.g., Davies et al., 1998; Petrides & Furnham, 2001, 2003, 2004). Saucier's "Mini-Markers" scale has a published average alpha reliability of 0.81 (Saucier, 1994) using a 9-point scale. The response format chosen for this study was a 5-point Likert scale to enable comparisons with existing data obtained from standard computer readable forms that only have five categories. The Likert scale categories were: *1-extremely inaccurate, 2-slightly inaccurate, 3-neither accurate nor inaccurate, 4-moderately accurate, 5-extremely accurate*. Reducing the number of Likert quantifiers from nine in Saucier's study, to five does not appear to have adversely affected the average reliability of 0.79 that was obtained in this study. Throughout this thesis, abbreviations of O, C, E, A, N will be used to specify the Big Five Factors (BFF) of Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (McCrae & John, 1992).

Word reasoning task (gc). The 23 items comprising this task originated from the Gf-Gc battery builder developed at Sydney University (Stankov, 1997; Stankov & Roberts, 2001). This test was included to provide a marker of crystallised intelligence, sometimes referred to as verbal intelligence, to establish convergent and divergent validity of the EI construct. It includes items of the following type: *"STATUE is to SHAPE as SONG is to …"*. Answers are selected from a list of four words, for example: *"BEAUTY, PIANO, TUNE, NOTE."* The participant is asked to choose the word that most closely completes the given analogy. This task serves as a marker for the broad ability of crystallised intelligence (Gc), and has an alpha reliability of 0.78 (Stankov & Roberts, 2001).

Tacit knowledge serial reaction time task. This test was included in the battery as it has been suggested that tacit knowledge or practical intelligence may be the underlying construct determining social or emotional intelligence (Sternberg & Hedlund, 2000). This task, a putative test of tacit learning ability (Berry, 1993), measures the participant's reaction time to pressing keys on

the computer keyboard when a stimulus appears on the screen in one of the four positions. The red-spot stimulus was one centimetre in diameter, appearing in one of four positions spaced six centimetres apart across the screen. This design was based on a serial reaction time (SRT) task as a measure of tacit learning ability (Nissen & Bullemer, 1987; Reber, 1993; Stadler & Frensch, 1998). When a specific stimulus appears in one of the four positions, the participant must press a corresponding key on the keyboard. The requisite keys are shown at the bottom of the screen and z, x, period, and / were chosen as they provide a comfortable resting position for fingers corresponding to position 1 to 4 respectively. The test has three phases: 10-practice trials, 40-random sequence trials, 96 repeated-sequence trials, and finally, 40 random-sequence trials. The repeated-sequence consists of eight digits corresponding to positions (13243142) repeated 12 times. Tacit learning occurs during the 96 repeated-sequence trials and is measured as the difference between average random-sequence response time and the average repeated-sequence response time.

Motivation. The motivation items were gathered from a variety of public-domain sources and reworded to elicit a response in the first person. Motivation is considered to be a component of EI in tests such as the Emotional Competence Inventory (ECI) (Boyatzis, Goleman & Rhee, 2000), and the EQ-i (Bar-On, 1997). Sample items from the 16-item Motivation questionnaire included in this study are:

"I think about running a company"

"I tell people what to do"

"I think about retiring early with millions in the bank"

Impulse Control. Impulse control is considered to be a component of EI (e.g., Goleman, 1995; Boyatzis et al., 2000; Bar-On, 2000). The 16 items in the Impulsivity scale were taken from the impulsivity scale in the Multidimensional Personality Questionnaire (MPQ) (Tellegen, 1982, 2002). Examples of items are:

“I often act without thinking”

“I am more likely to be fast and careless rather than slow and plodding”

“I often act on the spur of the moment”

Empathy. Empathy is assumed to be a component of trait or mixed EI theory (e.g., Goleman, 1995; Bar-On, 1997, 2000). The Empathy section of this battery contains 13 selected items from the *Interpersonal Reactivity Index,* (Davis 1980, 1983, 1991). The response format is a 5-point Likert scale with the following extremities: 1-*Strongly Disagree*, to 5-*Strongly agree*. Examples of some reverse-coded items are:

“I sometimes find it difficult to see things from the other person's point of view”

“I don't feel very sorry for other people when they are having problems”

“Other people's misfortunes do not usually disturb me”

Mach IV. Machiavellianism concerns the adaptive advantages of manipulative social behaviour. A high score on the Mach scale is a potential indicator of ability to read other people's emotions and use this information to manipulate them (Christie & Geis, 1970). As Machiavellianism does not correlate with general intelligence and does not consistently lead to real-world success (Christie & Geis, 1970; Wilson, Near & Miller, 1996) it was considered informative to see if it occupied the same psychological space as EI. The inclusion of this test was to determine if EI had any relationship with Machiavellianism in terms of emotion recognition and manipulation of others (e.g., Griffiths, 2001; Cialdini, 2001). The Mach IV test is a widely publicised and documented tool for measuring this construct, and typically has an alpha reliability of 0.73 (Zook & Sipps, 1985).

EI Tests used in this study

In addition to a wide variety of personality, cognitive and general trait tests listed above, two EI tests and a behavioural evaluation of EI was administered to the 115 first-year psychology participants.

The TEIQue-SF. The Trait Emotional Intelligence Questionnaire – Short Form (TEIQue-SF) is a 30-item form is designed to efficiently measure global Trait-EI (Petrides, Perez, & Furnham, 2003, 2004). For simplicity, the TEIQue-SF will be referred to as the TEIQue in this document. The TEIQue is freely available for researchers; factor structures, SPSS syntax, and norms were generously provided by K. V. Petrides for use in this study (personal communication, March 27th, 2006). The TEIQue provides reliable global trait EI scores that correlate meaningfully with a wide range of diverse criteria, including coping styles, life satisfaction, personality disorders, perceived job control, and job satisfaction (Petrides & Furnham, 2003, 2007). Items are measured on 7-point Likert scales with only the extremes labelled: *1-Completely disagree, and 7-Completely Agree*. The TEIQue differs from the factor structure of the ECA and includes the following four factors: Well-being, Self-control, Emotionality, and Sociability.

Sample *Well-being* factor items:

"On the whole I am pleased with my life"

"I generally believe that things will work out fine in my life"

Self-control:

"On the whole, I am able to deal with stress"

"Generally, I am able to adapt to new environments"

Emotionality:

"Expressing emotions with words is not a problem for me"

"I often pause and think about my feelings"

Sociability:

"I can deal effectively with people"

"I would describe myself as a good negotiator"

The Emotional Competence Assessment (ECA). The ECA was discussed in an earlier chapter, and has been an ongoing project at Macquarie University since 2003 (Reid, 2004, 2005, 2006). Norms and factor structure have been confirmed in a variety of studies with over 1500 participants. The ECA is a trait EI instrument based on the widely-accepted four-factor construct of emotional intelligence (e.g., Goleman et al., 2002; Bradberry & Greaves, 2004). The ECA has a 4–factor structure and utilises a 5-point Likert scale ranging from: *1-Never*, to *5-Always*. Bradberry and Greaves (2004) claim that this 4-factor structure utilised in their EIA test explained twice as much variance in job performance of leaders as the MSCEIT (Mayer et al., 2002c). The four factors in the ECA are: Self Awareness, Self management, Group awareness, and Group management.

Sample *Self-awareness* (SA) factor items are:

"I am aware of my feelings as they arise"

"I know the reason why I am feeling the way that I do"

Self-management (SM):

"If someone hurts me I get over it quickly"

"I enjoy revealing my feelings to friends"

Group-awareness (GA):

"I can detect other people's emotions"

"I sense when someone misreads the meaning of a group"

Group-management (GM):

"I am able to influence others"

"I am enlisted to resolve conflicts"

Brief Evaluation of Behaviour Interview (BEB). This measure was derived from a behavioural observation of the participant's personality style demonstrated in a brief (5-minute) videotaped interview. The design was modelled on Ford and Tisak (1983), however, instead of a single judge rating the behavioural characteristics there were four independent judges, who viewed

and coded the personality characteristics displayed by each participant in the 5-minute video clips. All four judges viewed all 115 participant video clips and rated behaviours on a ten-point scale. The participants were rated on eight measures: affect, smiling, responsiveness, fidgeting, posture, eye-contact, warmth, and confidence. These behaviours were selected from those used in previous studies (e.g., Ford & Tisak, 1983; Fletcher & Fitness, 1990; Funder & Colvin, 1993; Funder & Sneed, 1993). The judges' scores on a scale of 1 to 10 for each behaviour were bounded by "poor or unresponsive" at the low end and "very competent or appropriate" at the high end. The scoring was based on "appropriate" behaviour for the interview, for example, a high score would be given for smiling at the appropriate times, or demonstrating appropriate affect, without being excessively emotional or devoid of emotion. Eye contact and posture were also rated as appropriate or otherwise to an interview, while the amount of fidgeting was reverse scored as it represented inappropriate behaviour for an interview situation. The judges viewed the video clips once to establish their own personal norms and rating standards, then a second time to give the actual rating.

In addition to the eight measures listed above, the judges were also required to provide a total overall score of maturity and social skill for each participant. For this overall score the judges were directed to assume that each participant had applied for a job in a customer-facing role such as retail sales, or insurance claims processing, and that maturity and social skills were required for the job. The judges made this overall assessment of the participant's maturity and social competence at the end of the interview, after considering the participant's ratings on the eight behaviours listed above,. This represented a behaviourally-based EI score observed in the five-minute interview. This overall score is referred to as the brief evaluation of behaviour (BEB) and plays a similar role to the "Interview" score in the Ford and Tisak (1983) study.

During the private videotaped interview a fixed script of open questions was directed to the participant. These questions were designed to evoke an emotional response from the participant as most people find them difficult and personal at times. The questions used in the videotaped interview are listed below:

1. What do you want to be doing, or where do you want to be in 5-years time?
2. Do you see yourself as a leader, team player, nurturer or independent?
3. How do you tune in to the needs and feelings of your close friends?
4. How do you express warmth and concern for your friends?
5. What areas have you been a leader in?
6. How do you get people to do what you want them to do?
7. What is your birth order?
8. What do you think the role of the (substitute birth order from Q.7.) child is?

9. What was your most salient personality attribute at high school?
10. Where would you have come in a popularity poll of the class?
11. Do you believe that people have a specific purpose in life?
12. What would you like to be remembered for?

Academic performance criterion. The academic performance score for each participant was obtained from the Psychology 104 (PSY104) coordinator at the end of semester one, 2006, when results were available for publication. The PSY104 assessment included the following: a quiz consisting of 20 multiple choice questions in week 4 (10%); producing an annotated list of references to be used in essay (3%); a 5-page essay (22%); 4 hours of research participation (5%); and, a 2-hour final examination consisting of 120 multiple choice questions (60%). The academic score for each student was a scaled result out of 100% maximum.

Ethics

All procedures adopted in the study were approved by the Human Ethics Committee of Macquarie University, NSW, Australia.

Apparatus

All tasks for this study were computer-based using Dell, IBM-compatible computers with 2.8GHz clock speed, running Windows™ XP. All computer-based tasks were developed by the author to run on IBM-compatible personal computers using Macromedia Authorware™ 7 software.

Test Procedure

Participants took the test in a quiet, computer laboratory environment. The battery of tests was designed to take 45-minutes to allow for the brief videotaped EI interview and instructions within the one-hour period. Most participants completed within the hour. The experimenter explained the first item of each test in the sequence to ensure familiarity and understanding and remained with the participants for the entire test to ensure quality data, and to field queries as they arose. At the completion of the test battery, each participant was interviewed briefly and thanked for their participation.

Results

Descriptive Statistics

Table 10 lists the tests used in this study, with means, standard deviations and Cronbach's coefficient alpha internal reliability achieved for each test in this study.

Table 10

Descriptive statistics for all participants in the battery of tests (n=115) with internal reliability, coefficient alpha. Female and male scores with p-value of differences are also shown.

Test	All participants			Females		Males		*p-*
	M	*SD*	*Alpha*	*M*	*SD*	*M*	*SD*	*value*
1. Age of participants	22.6	7.55	-	24.1	8.32	22.5	4.65	*ns*
2. Facial affect recognition	0.80	0.09	.71	.80	.09	0.78	0.09	*ns*
3. Tracing score (total errors)	163	103	.88	163	109	164	88	*ns*
4. Personality questionnaire:								*ns*
Openness	30.9	4.6	.78	30.7	4.77	31.6	3.89	*ns*
Conscientiousness	28.4	5.1	.78	28.1	4.97	29.2	5.58	*ns*
Extraversion	26.9	5.4	.81	27.2	5.37	26.4	5.56	*ns*
Agreeableness	32.6	4.4	.81	33.4	3.94	30.6	5.01	<.01
Neuroticism	20.4	4.8	.79	20.9	4.57	18.9	5.20	<.05
5. Word reasoning task	0.77	.15	.71	0.78	.14	0.74	.17	*ns*
6. Tacit learning task	0.04	.04	.93	0.04	.04	0.04	.03	*ns*
7. Motivation questionnaire	3.36	.48	.69	3.33	.49	3.44	.45	*ns*
8. Impulse control	3.33	.57	.84	3.33	.58	3.34	.54	*ns*
9. Empathy scale	3.76	.61	.77	3.80	.59	3.63	.66	*ns*
10. Mach IV scale	2.64	.67	.72	2.61	.66	2.73	.70	*ns*
11. Trait EIQ (TEIQue)	5.07	.65	.86	5.02	.65	5.21	.65	*ns*
12. Emotional competence assessment (ECA)	3.53	.32	.89	3.57	.36	3.73	.38	<.05
13. Interview score (BEB)	0.00	.85	.90	.13	.85	-.34	.79	<.01
14. PSY104 Semester 1 psychology mark	61.9	13.4	-	61.70	14.5	62.4	10.0	*ns*

Behavioural observations

The behavioural interview was conducted after the above battery of tests was completed by each participant. The interview was videotaped and, at a later date, the four judges rated each participant on their overall social intelligence and maturity score. This behavioural rating is used throughout this study as an operationalised (real-world) measure of EI, and for convenience, will be referred to as the brief evaluation of behaviour (BEB) throughout this thesis. All four judges independently viewed and rated the 115 participants to achieve the highest possible consistency in the scoring of the BEB score. The four judges' scoring was converted to standard scores (z-scores), and intraclass correlation coefficient statistics were obtained for "consistency" of the judges' scoring using the "Two-Way Random" model in SPSS (Strube, 2002). Thus, this brief evaluation of behaviour (BEB) score is represented with a mean of zero and a standard deviation of one.

After calculating the reliability of the BEB score, the cases where the judges had given disparate scores (identified by an SD greater than one between judges' ratings) were reviewed by the group of four judges, and disagreements in the scoring were resolved through discussion. The

changes to individual judge's scores were minimal, affecting less than 5% of the overall scores. The initial reliability before adjustments was 0.89, and the final reliability after discussion and adjustment was 0.90. The judges' scores were then averaged to provide a composite overall score for BEB. The intraclass correlation reliability of 0.89 can be considered high by psychometric standards (Murphy & Davidshofer, 1998) even before discussion and adjustments were made to raise it to 0.90. The high reliability of 0.89 provides evidence that the judges consistently rated the participant's behaviour, even before discussion and adjustment, and that further statistical analysis may proceed with confidence.

Correlation Matrices

Table 11 demonstrates the positive manifold between EI test factors for both the TEIQue and the ECA that is necessary, but not sufficient, evidence for a unitary EI intelligence factor. Table 12 shows the intercorrelations between all tests and criteria used in this study.

Table 11

Correlation matrix of the EI tests included in this study. Note that all EI tests and the total EI scores are positively correlated with each other.

EI Scales and subscales	1	2	3	4	5	6	7	8	9
1. Total TEIQue score	1								
2. Total ECA score	.66**	1							
3. TEIQue Well-being	.84**	.38**	1						
4. TEIQue Self-control	.76**	.39**	.61**	1					
5. TEIQue Emotionality	.76**	.57**	.53**	.40**	1				
6. TEIQue Sociability	.59**	.65**	.31**	.24*	.29**	1			
7. ECA Self-awareness	.75**	.61**	.65**	.61**	.59**	.35**	1		
8. ECA Self-management	.44**	.67**	.25**	.16	.53**	.38**	.26**	1	
9. ECA Group Awareness	.34**	.71**	.12	.18	.29**	.46**	.18	.28**	1
10. ECA Group Management	.30**	.79**	.04	.10	.21*	.61**	.21*	.39**	.57**

Note: * = correlation is significant at the $p < .05$ level, ** = correlation is significant at the $p < .01$ level

Table 12

Correlation matrix of test scores in this study, with the two criteria: the BEB interview score and the PSY104 academic results score shown as the first scores on the list.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. BEB	1														
2. PSY104	-.16	1													
3. Faces	.12	.24**	1												
4. Verbal Reasoning	-.06	.39**	.38**	1											
5. Tracing	.00	.08	.18	.30**	1										
6. Motivation	.22*	-.34**	-.16	-.29**	-.26**	1									
7. Impulse control	-.17	.28**	.02	.22*	.09	-.04	1								
8. Empathy	.00	.07	.32**	.23*	.09	-.26**	.17	1							
9. Mach IV	.03	-.15	-.16	-.13	-.14	.23*	-.19*	-.39**	1						
10. Openness	-.01	.06	.18	.16	-.14	.29**	.20*	.26**	-.06	1					
11. Conscientiousness	-.12	.09	-.21*	-.19*	-.14	.31**	.56**	-.13	-.11	.04	1				
12. Extraversion	.37**	-.09	.05	-.12	-.05	.42**	-.10	-.04	-.02	.19*	.15	1			
13. Agreeableness	.13	-.11	.16	.16	.12	-.08	.25**	.56**	-.29**	.24*	.08	.02	1		
14. Neuroticism	-.02	-.11	-.02	-.11	-.25**	.09	-.15	-.20*	.43**	-.08	-.16	-.30**	-.15	1	
15. Total TEIQue score	.21*	.18	.17	.10	.13	.22*	.14	.27**	-.38**	.26**	.12	.44**	.26**	-.52**	1
16. Total ECA score	.12	-.03	.22*	.04	-.11	.39**	.12	.26**	-.11	.45**	.08	.36**	.26**	-.16	.66**

Note: * = correlation is significant at the $p < .05$ level, ** = correlation is significant at the $p < .01$ level

Testing Hypothesis 1

H1: If EI is indeed a unified intelligence that can be represented by a single global factor, then correlation matrix patterns and structural equation models (SEM) will confirm this structure.

This hypothesis was examined from two angles to determine whether the empirical data collected in this study could be explained by a single latent EI factor that underlies all of the EI tests and subfactor scores: (A) Evidence of a positive manifold between tests and subfactors; and, (B) A structural equation model (SEM) providing confirmatory evidence of a unitary EI factor in all EI tests and subfactors.

(A) Evidence of a positive manifold indicating a unitary structure of EI. This evidence is based in the logic that Spearman (1904) used to propose a general factor of intelligence (g) from a variety of subscales that were found to correlate positively. The correlation matrix in Table 11 shows that all EI tests and total EI scores are positively correlated with all tests, total scores, and subscales. This demonstrates a "positive manifold" (Spearman, 1904) between all tests that can be put forward as necessary, but not sufficient, evidence that EI is a unitary structure.

(B) Structural equation modelling (SEM). If the empirical data obtained in this study can support a model whereby the subscales in the TEIQue and the ECA, both separately and combined, can be explained by a single latent factor of EI, then this can be offered as evidence for the existence of EI as a unitary factor, as proposed by EI theorists (e.g., Goleman, 1995; Mayer, Caruso & Salovey, 1993, 1999, 2000b, 2000c).

For this study, the cut-off values used to determine SEM model fit were taken from the AMOS 6.0 User's Guide (Arbuckle, 2005, 2005a) and consisted of:

1. The *chi-square* test of model fit (CMIN) expressed as a *p-value,* which contrary to normal statistical testing should be greater than 0.05 to indicate that the data fits the model. This test is susceptible to sample size confounds; smaller samples are likely to demonstrate a better fit so the following three indices are preferred.
2. A *chi-square/df* (CMIN/DF) ratio of less than 3 indicates an acceptable fit of the model to the data.
3. A Tucker-Lewis Index (TLI) of greater than 0.95 is considered an acceptable fit, while values close to 1 indicate a very good fit of the model to the data.
4. A root mean square error of approximation (RMSEA) of less than 0.05 indicates a close fit, <0.08 indicates a reasonable fit, with >0.10 a poor fit of the model to the data.

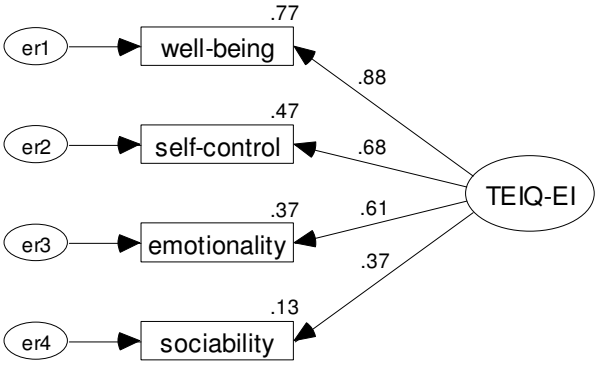


Figure 3

Structural equation model showing the four subfactors of the TEIQue. Model fit statistics were within accepted norms and thus confirm this model as a valid representation of the data (*p=0.56*, *CMIN/DF=0.58, TLI=1.03, RMSEA=0.00).*

The structural equation model in Figure 3 indicates that the four subfactors of the TEIQue EI test can be represented by a single latent EI factor (TEIQ-EI). The low loading on the sociability subscale (0.37) and low variance explained (0.13) by the latent EI factor suggests that the data may be better described by a 2-factor model.

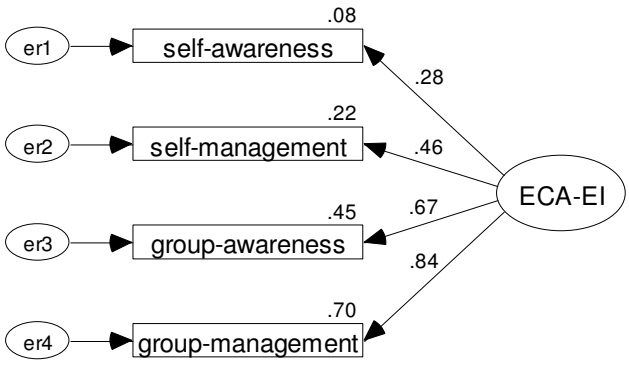


Figure 4

Structural equation model showing the four subfactors of the ECA. Model fit statistics were within accepted norms and thus confirm this model as a valid representation of the data (*p=0.23*, *CMIN/DF=1.45, TLI=0.96, RMSEA=0.06)*.

The structural equation model in Figure 4 indicates that the four subfactors of the ECA EI test can be represented by a single latent EI factor (ECA-EI). The low loading on the self-awareness subscale (0.28) and low variance explained (0.08) by the latent EI factor suggests that the data may be better described by a 2-factor model.

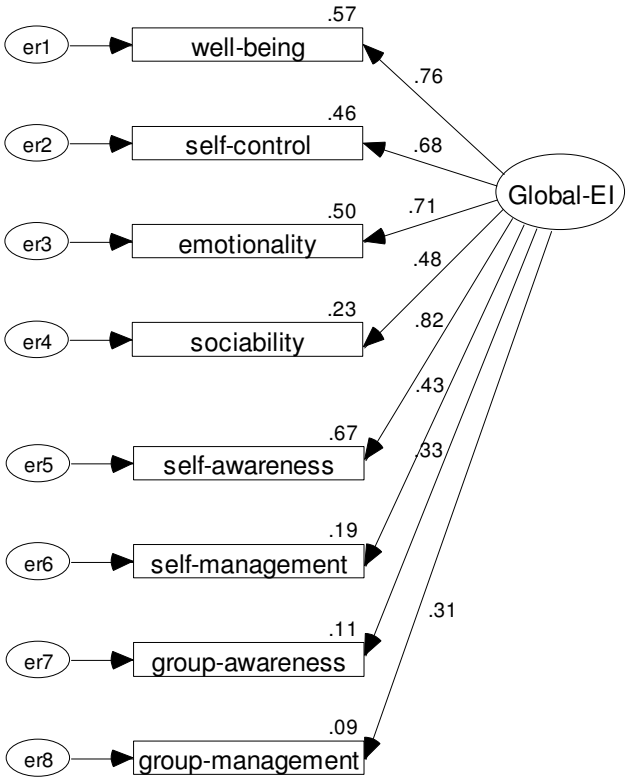


Figure 5

Structural equation model including the eight subfactors of the ECA and the TEIQue. Model fit statistics were outside accepted limits thus disconfirming this model as a potential representation of the data (*p=0.000*, *CMIN/DF=7.23, TLI=0.49, RMSEA=0.23)*.

The structural equation model shown in Figure 5 confirms that the eight subfactors of the ECA and the TEIQue EI test cannot be represented by a single latent EI factor (Global-EI) because all fit indices exceed acceptable limits.

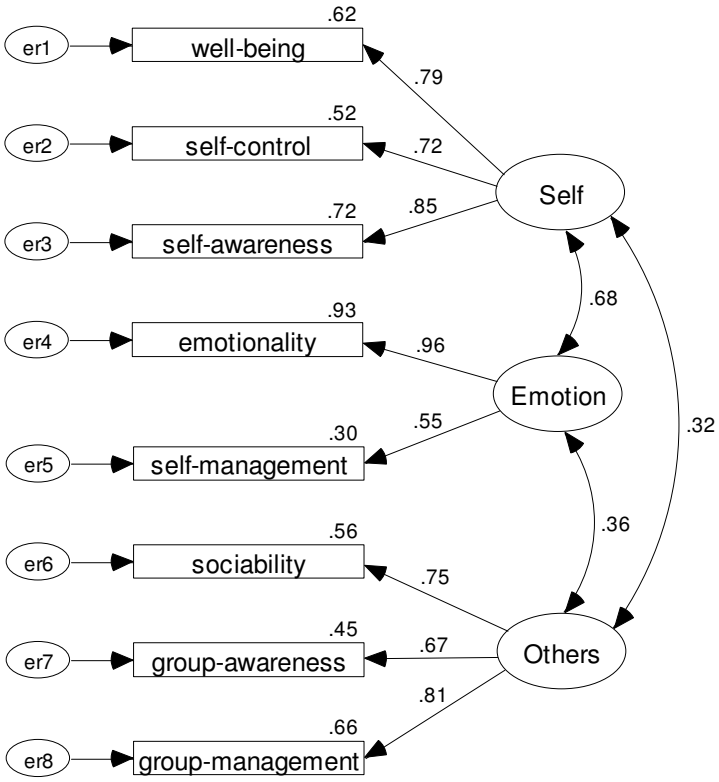


Figure 6

A three-factor model using the eight subfactors of the ECA and the TEIQue. (*p=0.000, CMIN/DF=2.59, TLI=0.87, RMSEA=0.12).*

A three-factor model was established using exploratory factor analysis in SPSS, then confirmed using AMOS SEM. This model fits the data better than the single-factor model in Figure 5; however, the fit statistics are still outside acceptable limits. This model, with some refinement and tuning, may turn out to be a useful representation of the data; however, it clearly does not fall within acceptable fit indices at this stage. Even if the model could be refined to the point where it fell within acceptable fit indices, it still would not provide evidence for a global EI factor as it contains three latent factors.

To summarise these findings, the models shown above demonstrate that while the TEIQue and the ECA, separately, can be verified as constructs with a single global EI factor, a single factor solution is not supported when the two EI test subfactors are combined in one model. Hypothesis one is therefore rejected; the data gathered from the EI tests used in this study cannot be explained by a single global EI factor.

Testing Hypothesis 2

H2: If EI is a new and useful psychological construct, then it will explain significant incremental variance above and beyond existing personality measures, when predicting emotionally intelligent behaviour measured in a behavioural interview (BEB).

Regression analysis – general predictive properties of the EI tests. Regression analysis was used to determine which computer-based tests best predict the emotionally intelligent behavioural criterion obtained from the 5-minute videotaped interview (BEB). During the interview, participants were rated on their overall maturity and positive interview attributes to provide a measure of emotionally intelligent behaviour.

To determine the best predictors of the BEB interview score criterion, all test results were subjected to stepwise and backward regression analysis utilising all predictors obtained from the tests to ensure that no significant models had escaped attention. Although stepwise regression provides a list of models that could apply to the data, only the models reaching statistical significance, or of value for our discussion, will be discussed below. The predictor labelled BFFEI was constructed by adding the four "positive" factors of the BFF model (O, C, E, A) and subtracting the Neuroticism score as proposed by McCrae (2000). The predictor labelled "EI Subfactors" consists of the eight TEIQue and ECA subfactors entered individually rather than as total EI scores. These subfactors are: well-being, self-control, emotionality, and sociability (TEIQue), and, self-awareness, self-management, group-awareness, and group-management (ECA).

Table 13

Summary of variance explained by EI and personality tests.

Model	Predictors for BEB (the EI Interview criterion)	Variance explained	Model significance *p-value*
1.	TEIQue	5%	0.02
2.	ECA - Emotional Competence Assessment	2%	0.19
3.	TEIQue, ECA	5%	0.07
4.	BFFEI	2%	0.11
5.	-O, -C, E, A, N (BFF)	21%	0.001
6.	EI-subfactors from the TEIQue and ECA	20%	0.002

Each model in Table 13 represents the predictive properties of EI or personality tests for the BEB criterion in terms of variance explained, and model significance. Models assume a positive increase in the BEB criterion with increases in the individual predictors, unless shown with a minus sign indicating a negative relationship.

Model 1 in the regression analysis shown in Table 13 indicates that the TEIQue EI test explains 5% of the variance in the BEB scores. Model 2 indicates that the ECA EI test, based on the Goleman (2002) model of EI, only explains 2% of the variance in the BEB criterion score. Model 3 consisting of the TEIQue and the ECA global scores explains 5% of the variance. Model 4 which is named "Big Five Factor EI" (BFFEI) score shows a variance of only 2% explained when the BFF personality scores are constrained according to the model suggested by McCrae, that is, that O, C, E and A, should have positive, while N should have negative relations with EI scores (McCrae, 2000). The predictive power of the constrained BFF construct (BFFEI) is thus similar in magnitude the predictive power of the global EI scores for the TEIQue and the ECA (models 1 & 2 respectively).

The Big Five Factors of personality when unconstrained in a regression model explain 21% of the variance in the BEB criterion (model 5). This is an order of magnitude larger than when the BFF factors are constrained to McCrae's model (model 4), indicating that unconstrained subfactors, be they EI or BFF, have much greater predictive power. Finally, model 6 shows that 20% of the variance in the BEB criterion can be explained by the EI subfactors. This value is similar to the explanatory power of the BFF personality factors shown in model 5.

Incremental validity analysis for the BEB criterion. The following analysis was conducted using hierarchical linear regression to assess the incremental validity of the EI tests above and beyond the five personality factors of the BFF model (Costa & McCrae, 1985, 1992; McCrae & John, 1992). While it is acknowledged that some of the predictors do reach significance (e.g., O, C, E, A, N), the purpose of including these covariates is to control for BFF personality to establish whether EI adds incremental validity.

Table 14

Summary of variance explained in the BEB criterion by various combinations of EI and personality tests.

Model	Predictors for BEB (EI interview score)	Std. Beta Estimates	Parameter p – value	Incremental variance over model 1	Total variance explained R^2
1.	Openness (O)	-.12	.17		21%
	Conscientiousness (C)	-.18	.04		
	Extraversion (E)	.45	<.001		
	Agreeableness (A)	.18	.05		
	Neuroticism (N)	.10	.29		
2.	-O, -C, E, A, N	-	-	1%	22%
	TEIQue	.13	.26		
3.	-O, -C, E, A, N	-	-	1%	22%
	ECA	.00	.98		
4.	-O, -C, E, A, N			2%	23%
	TEIQue	.20	.15		
	ECA	-.11	.38		
5.	1. -O, -C, E, A, -N	-	-	11%	32%
	2. ECA – Self-awareness	-.40	.004		
	3. ECA – Self-management	.16	.21		
	4. ECA – Group-awareness	-.27	.02		
	5. ECA – Group-management	.22	.10		
	6. TEIQue – Well-being	.21	.10		
	7. TEIQue – Self-control	.11	.43		
	8. TEIQue - Emotionality	.23	.08		
	9. TEIQue - Sociability	-.07	.59		
6.	1. -O, -C, E, A, -N	-	-	3%	24%
	2. Motivation	-.20	.06		

Table 14 indicates the total and incremental variance explained by various groupings of the predictors for BEB, the criterion. All models are significant with overall fit statistics less than p=.001.

Model 1 is the baseline variance explained by the BFF personality factors, to which the various EI tests are incrementally added. Referring back to the zero-order correlation matrix shown in Table 12, Extraversion, Motivation, and the TEIQue scores are significantly correlated ($p < 0.05$) with the BEB criterion. This indicates that 14%, 5%, and 5% of the variance is explained by these

predictors respectively. Overall, Extraversion appears to be the single best predictor for the BEB criterion in this battery of tests. Be that as it may, all BFF personality factors are usually included when examining incremental validity of EI; in this case they explain 21% of the variance as shown in model 1 of Table 14. Note that Openness and Conscientiousness have negative beta weights indicating that with all other factors controlled for, an increase in Openness or Conscientiousness is associated with a decrease in the BEB EI criterion.

The 16-item Motivation scale explained an additional 3% of the variance in the BEB criterion after controlling for personality with the BFF factors (model 6). Mixed EI tests often include Motivation as an EI factor (e.g., Goleman, 1995; Boyatzis et al., 2000; Bar-On, 1997, 2000); this will be examined further in the discussion.

Since the total EI total scores using the TEIQue, the ECA, separately and combined explained insignificant incremental variance (2%) with the BFF controlled for (see models 2, 3 & 4), it was decided to check if the unconstrained EI subfactors would increase the predictive power of the EI tests. After adding the TEIQue and the ECA subfactors as predictors for the BEB score, a significant 11% of incremental variance ($p = 0.05$) was explained (model 5). The TEIQue subfactors, Well-being and Emotionality, and the ECA subfactors, Self-awareness (SA), Group-awareness (GA), and Group Management (GM) were the most significant predictors for the BEB criterion. The EI factors explain 11% of additional variance (model 5 over model 1 in Table 14).

These results demonstrate support for hypothesis two; EI subfactors explain significant incremental variance above and beyond existing personality (BFF) factors. However, hypothesis two is not supported when total (global) EI scores for the TEIQue and ECA are used, that is, they do not explain significant incremental variance. This has implications for hypothesis one. If EI is unitary construct, then total EI scores for the TEIQue and ECA should demonstrate similar incremental validity to that which the EI subfactors demonstrate. Thus, results from this analysis provides further evidence against hypothesis one; EI cannot be represented by a single global factor for the TEIQue and the ECA.

Testing Hypothesis 3

H3: If EI is a new and useful psychological construct, then it will explain significant incremental variance over existing intelligence and personality measures, when predicting first-semester psychology academic results (PSY104).

Incremental validity analysis for the PSY104 criterion. A regression analysis was conducted to determine whether EI has anything to add as a predictor for a student's performance in first year psychology, using the PSY104 mark as the criterion. The test battery results were added as predictors in the regression analysis, and those with highest predictive ability (lowest *p-value*) were

grouped to form a single model, as listed below in Table 15. Stepwise and Backward regression analysis utilising all predictors obtained from the test battery was also conducted to double-check that no significant models had escaped attention.

Table 15

Summary of variance explained in the PSY104 criterion by various combinations of cognitive, EI, and personality tests.

Model	Predictors for first semester Psychology I (PSY104) results	Total variance explained	Incremental variance explained over model 2	Incremental variance significance *p*-value
1.	Word	15%	-	0.0001
2.	Word, O, C, -E, -A, -N	23%	-	0.08
3.	Word, O, C, -E, -A, N, TEIQue	27%	4%	0.02
4.	Word, O, C, -E, -A, N, ECA	23%	<1%	0.95
5.	Word, O, C, -E, -A, N, TEIQue, ECA	29%	6%	0.001
6.	Word, O, C, -E, -A, N, plus eight EI-subfactors	34%	11%	0.04
7.	Word, O, C, -E, -A, N, eight EI-subfactors, plus Motivation	45%	22%	0.0001

Each line in Table 15 represents a separate model, with beta coefficients indicated by a minus sign where negative. Most models, except 2 and 4, demonstrate significant incremental variance ($p < 0.05$) over the Word IQ marker.

The correlation matrix shown in Table 12 indicates that the Word Reasoning Task (Word), a measure of general intelligence, predominantly crystallised intelligence (g_c), was the best predictor of the PSY104 results with a correlation of 0.39 representing 15% if the variance in the PSY104 criterion. It was expected that the Tracing task would tap fluid intelligence, as is had a significant correlation with the Word task ($r = .30$); however, its correlation with the PSY104 scores was negligible ($r = .08$), so it was omitted from further analysis.

Model 2 of the regression analysis shown in Table 15 above indicates that the Word task and the BFF personality factors explain 23% of the variance in the PSY104 scores. The Word task was the most significant correlate of success in PSY104 due to its primary role in the general intelligence model (Carroll, 1993). Of interest, Extraversion and Agreeableness had negative beta coefficients (noted by minus signs in the predictors column above), indicating that with other predictors controlled for, a decrease in Agreeableness or Extraversion is associated with higher scores in PSY104. Neuroticism had a negative beta weight indicating that, as expected, a decrease

in Neuroticism was associated with an increase in PSY104 scores. Although it is not typical to include personality factors in academic intelligence testing scenarios, it appears that the BFF personality construct has a contribution to make as a predictor of PSY104 scores as it explains 8% of incremental variance over crystallised intelligence (model 2 minus model 1).

Having covered the mainstream predictors of the PSY104 results, general intelligence, and personality, EI was tested for incremental predictive validity for academic success (e.g., Schutte, 1998; Petrides et al., 2004). Model 3 in Table 15 above indicates 4% additional variance explained by the global TEIQue score when added to the predictors in model 2 (crystallised intelligence and personality). The addition of the global ECA score as a predictor as shown in Model 4 did not explain any worthwhile incremental variance beyond model 2, however when added to the TEIQue as a predictor it increased the incremental variance explained by EI up to 6% (model 5).

Finally, as shown in model 6, all eight EI subfactors were added as predictors for the PSY104 criterion which explained a significant incremental variance of 11%. This result supports hypothesis three, that EI subfactors explain significant incremental variance over existing intelligence and personality measures, when predicting first-semester psychology academic results (PSY104).

Returning to hypothesis one; the variance explained when using EI subfactors (11%) is almost twice that of the next best model using the global TEIQue and ECA EI scores (Model 5) that explains 6% of incremental variance. This is further evidence to reject hypothesis one on the basis that EI is only half as predictive when represented by global factors.

The addition of the short Motivation scale increased the variance explained by an additional 11% (model 7). This suggests that the EI tests used in this study did not explain all the variance available beyond personality and IQ.

Finally, there were a number of tests in this battery that did not achieve significance as predictors of the BEB or the PSY104 scores, and thus have not been discussed in detail. These tests include, Machiavellian style (MACH IV), Empathy, Tracing as a measure of fluid intelligence (g_f), Tacit learning ability, and Impulse control. Impulse control had significant correlations with PSY104 ($p < .05$) in the zero-order correlation matrix (Table 12), but lost predictive validity after controlling for BFF personality, perhaps due to its high correlation with BFF Conscientiousness ($r = 0.56$).

Factor analysis for this study

The putative tests of emotional intelligence (EI) - the TEIQue and ECA - plus the general (academic) intelligence markers were subjected to a principal components analysis with orthogonal (varimax) rotation. This structure of this table is similar to that found in the Ford and Tisak study (1983) which provided evidence to conclude that they had found a social intelligence (SI) factor

that was divergent from academic intelligence. The results shown in Table 16 indicate two factors: an academic intelligence factor, and an EI factor. Communalities (h^2) representing the sum of all the common factor variance, are greater than 0.30 and thus adequate for interpretation of the solution (Child, 1976; Harman, 1976). The Tracing communality was a bit low, but close enough to 0.30 to accept the results.

Table 16

Principal Components analysis for the current study indicating a 2-factor solution, general intelligence, and EI.

Varimax rotated factors			
Test	h^2	1	2
PSY104	.48	.69	
Faces	.45	.59	
Tracing	.26	.51	
Word	.64	.80	
TEIQue	.64		.77
ECA	.69		.84
BEB (Interview)	.32		.54

Table 16 shows orthogonally rotated factors from a principal components analysis demonstrating two factors, academic intelligence (1) and emotional intelligence (2). All tests demonstrate a single and substantial loading on one factor or the other using a recommended cut-off of 0.30 for loadings (Child, 2006). The TEIQue, the ECA, and the interview (BEB) loadings cluster to form an EI factor divergent from the academic or general intelligence, factor 1. The communality of the Tracing task is low considering the recommended cutoff of 0.30 (Child, 2006), suggesting that it may be assessing some other unknown factor in addition to general intelligence (*g*).

The first-semester psychology I results (PSY104), the facial emotion recognition task (Faces), Tracing, and the Word reasoning task cluster to form factor 1 representing a latent factor of general intelligence. Acuity in recognising emotions in faces has been proposed as part of EI (Mayer et al., 2002c); however, this factor analytic evidence shows that it is related to general intelligence rather than EI.

Adding personality to the analysis. So far, the Ford and Tisak (1983) method of factor analysis was carried out using general intelligence markers and potential EI constructs, leading to two divergent factors. However, it is probable that personality theory (Costa & McCrae, 1985, 1992) had not reached the ascendant position in social psychology research in 1983 that it holds today; thus Ford and Tisak did not control for personality in their study. When the personality

factors are included in the factor analysis they appear to cluster with the EI factor. Thus, the research question becomes whether the EI factor is distinct and divergent, or just part of the BFF personality construct.

Table 17

Personality added to the EI variables in the Principal Components analysis.

Varimax rotated factors			
Test	h^2	1	2
PSY104	.34		.59
Faces	.44		.62
Tracing	.27		.52
Word	.65		.80
TEIQue	.67	.81	
ECA	.59	.77	
BEB	.18	.40	
Openness	.25	.49	
Conscientiousness	.14	-.31	
Extraversion	.52	.67	
Agreeableness	.25	.47	
Neuroticism	.25	-.48	

Table 17 indicates that when personality factors are added to the existing EI measures in the analysis, personality tends to cluster with the EI factors. This analysis has been constrained to a 2-factor solution to demonstrate the EI/personality versus general intelligence dichotomy, thus communalities (h^2) are low in some cases (< 0.30). In summary, Table 17 indicates that the major tests used in this study fall into two psychometric categories: the general intelligence domain (IQ), and the EI-personality domain. This confirms that the tests data conform to the same patterns reported by Ford and Tisak (1983). Therefore, following their methodology, it could be concluded that factor 1 represents an independent EI factor. However, factor 1 may alternatively represent the personality domain; a hypothesis that Ford and Tisak did not explore.

In summary, EI as represented by the TEIQue and ECA tests did not exhibit a unified structure when examined using SEM. Furthermore, the EI subfactors explained almost twice the variance that the global EI scores explained. Thus, hypothesis one (H1) was rejected on two separate pieces of evidence. In terms of predicting the behavioural interview score (BEB), EI subfactors explained an incremental variance over personality factors of 11%, thus providing support for hypothesis two (H2). When predicting academic achievement as measured by PSY104 criterion, EI subfactors were able explain an incremental variance of 11% over IQ and personality scores. This provides sufficient evidence to support hypothesis three (H3).

Discussion

The overall aim of this study was to replicate and expand the Ford and Tisak (1983) study using contemporary predictor tests in the updated context of EI, rather than Social Intelligence. Formal hypotheses were developed to: i) establish whether EI is a unified intelligence that can be represented by a single global factor (H1); ii) establish if putative tests of EI could predict a behavioural outcome representing EI, operationalised by the BEB interview (H2); and iii) determine if EI tests predict performance in first-year, first-semester psychology as measured by the PSY104 results (H3). In the following exposition, the results obtained in this current study are discussed in terms of the aims and hypotheses.

EI as a single factor (H1)

Structural Equation Modelling (SEM) demonstrated that the EI test subfactors could not be represented by a single higher-level EI factor, thus hypothesis one was rejected. Results indicate that the subfactors behaved as collection of loosely-related mixed tests (Ciarrochi & Godsell, 2006; Matthews et al., 2002; Mayer et al., 2000); a collection of unrelated non-intellective factors (Wechsler, 1940, 1950). Testing of hypothesis two found that the eight subfactors of EI predicted the behavioural criterion (BEB) significantly better than a single global EI factor. Incremental variance explained using both EI global test scores was an insignificant 2%, while the eight separate EI subfactors explained a significant 11% of the variance, when controlling for the BFF personality factors. Clearly, this five-fold increase in the variance explained by subfactors provides evidence for the non-additive structure of the EI subtests (e.g., Michell, 1990). The predictive power of the subfactors appear to cancel each other out when they are summed into a global EI score. This raises issues with the current simplistic practice of adding EI subfactors together, especially when they are almost orthogonal, as demonstrated in Figure 6, where the correlation between latent factors "Self" and "Others" is low ($r = 0.32$).

The EI literature has generally proceeded on the basis that EI can be represented by a single factor representing the various EI tests (e.g., Goleman, 1995; Mayer, Caruso & Salovey, 2002c). Others such as Bar-on (1997, 2000) have positioned EQ as non-intellective factors that influence intelligent behaviour; however, Bar-On still produces a total EQ score by summing the subfactors. The 15 subscales in the EQ-i have average intercorrelations between subscales of 0.50; however, with some scales having intercorrelations of 0.75, others must be quite low (Bar-On, 2000). Thus the validity of the general practice of adding together subscales that are essentially divergent, to produce a global EI (or EQ) score, must be questioned on the basis that it makes little psychometric sense to sum divergent constructs (Michell, 1990). To draw a parallel with personality measures, it is not considered appropriate to sum BFF personality factors to produce a total "personality" score.

In the same vein, it has been shown in this study that the EI constructs appear to lose their predictive properties when their sub-scores are simply summed into a composite global EI score.

EI as a predictor of maturity and superior interpersonal skills (H2)

The results show that hypothesis two was supported in regard to EI tests significantly predicting the behavioural outcome (BEB). The eight EI subfactors entered separately as predictors explained significant incremental variance in the BEB criterion above and beyond existing personality (BFF) factors. However, hypothesis two was not supported when global EI scores for the TEIQue and ECA were used as predictors. This is evidence for the rejection of hypothesis one: EI does not appear to be a unitary construct, on the basis that total global EI scores severely reduce the predictive power of the EI tests.

In terms of the other predictors, the finding that Extraversion was the pre-eminent predictor of emotionally and socially intelligent behaviour was unexpected considering that Goleman had proposed that EI was synonymous with good *character* (1995) and *maturity* (1998). While the behavioural criterion (BEB) had been designed to evaluate the emotional maturity and interpersonal skills of the participants, it transpired that the zero-order correlation was insignificant between the BEB criterion and the ECA, and low on the TEIQue EI test (see Table 12). The Extraversion factor was clearly the best predictor for the BEB criterion, although normally, *maturity* and *character* are not the primary defining characteristics for Extraversion (e.g., Costa & McCrae, 1985, 1992).

The results indicated that the EI tests, the TEIQue and the ECA each explained only 1% of incremental variance in the BEB score derived from the 5-minute interview (models 2 and 3 in Table 14). This was much smaller than expected, although the Ford and Tisak (1983) study never discussed the amount of variance explained by each test, and instead, relied on factor analytic evidence for SI. On the other hand, the Big Five personality factors (BFF) explained a much larger 21% of the variance of the BEB score as shown in model 1, Table 14. Overall, the EI test global scores were not very effective in terms of predicting the BEB criterion.

Above and beyond the EI tests, the short 15-item Motivation test developed for this study explained a 3% incremental variance over the BFF personality scores. In terms of efficiency and predictive power, the 16-item Motivation scale exceeded the EI tests which have 84 items combined. If Motivation is considered an EI factor (e.g., Goleman, 1995; Boyatzis et al., 2000; Bar-On, 2000) it can make a substantial contribution to the predictive validity of self-report EI or "mixed" EI tests (Mayer et al., 2000). Motivation was also a substantial predictor for the first-year psychology results criterion (PSY104) and is discussed further below.

Constrained EI factors with unconstrained BFF personality factors. Subfactors that conform to a theoretical model are constrained by the model. EI subtests are constrained if they are assumed to correlate positively with each other and are additive, thus generating a global EI score (e.g.,

Mayer et al., 2002; Bar-On, 1997; Goleman, 2002). On the other hand, the BFF personality scores are unconstrained and are not required, *a priori*, to have a positive or negative relationship with each other, or an outcome variable such as BEB. During regression, the BFF personality predictors are advantageously attuned to explaining the variance in an outcome variable such as BEB due the unconstrained regression beta weights. By contrast, the EI factors are constrained to being positive due to the EI model assuming a global EI factor made up of additive subfactors.

The BFF personality scores (Openness, Agreeableness, etc.) were the most predictive family of factors for the BEB (interview) criterion, explaining 21% of the variance (Table 14). When the TEIQue total EI score is added to the BFF predictors, the incremental variance explained only increases by one percent. Researchers in the past (e.g., Davies et al., 1998; Matthews et al., 2002, 2004) have proposed this as evidence that self-report EI tests are simply a proxy for personality because they add so little to the variance when personality is controlled for. The BFF factors will, however, tend to be ascendant if they are allowed to freely establish their optimal regression coefficients, without some sort of theoretical model specifying what BFF factors, and their directionalities are associated with emotionally intelligent behaviour. In regression analysis, using unconstrained BFF scores with a constrained global EI score is an "apples versus oranges" type of comparison, tending to advantage the unconstrained factors.

Robert McCrae (2000), co-author of the Big Five theory of personality (Costa & McCrae, 1985, 1992; McCrae & John, 1992) asserts that EI should be positively related with high scores in Extraversion, Openness, Agreeableness and Conscientiousness, and inversely related with Neuroticism (the O+C+E+A-N, or the "four minus one" model). Following McCrae's specification of a model I tested the composite "four minus one" model, the BFFEI score, by adding the positive factor scores and subtracting the negative factor, Neuroticism. This parallels the method used to produce total global EI scores in the TEIQue and ECA, based on the EI model.

The results shown in model 4 in Table 13 indicate that when the BFF factors are constrained with McCrae's (2000) model, the variance explained reduces dramatically to 2%. The unconstrained BFF model by contrast explains ten times as much variance, 21% (model 5). The negative beta weights on O and C in Model 4 are at odds with McCrae's assertion that all factors except N should be positively related to emotionally intelligent behaviour.

The process of regression achieves minimum error and maximum variance-explained by selecting optimum beta weights for the independent predictor variables for the criterion, in this case BEB. When the beta weights for the predictors are constrained, such as in the BFFEI model, the regression procedure only has one predictor that it can adjust to achieve the best fit, hence dramatically reducing the variance that can be explained by the model. Constraining predictors is

not the most effective way of using the independent measures at our disposal to optimise the variance-explained, when predicting emotionally intelligent behaviour.

With the EI factors constrained to a model, and on the other hand, the BFF personality factors being unconstrained, the BFF factors become paramount in explaining most of the variance, leaving little variance over for the single factor EI test to explain. When it comes to deciding whether EI tests have any predictive validity above and beyond the BFF factors, this approach leads to the erroneous conclusion that EI tests add very little. This is because it is based on an inequitable comparison; a category error in terms of comparisons.

Unconstrained EI factors with unconstrained BFF factors. When the subfactors of the TEIQue and ECA were used as predictors in a regression analysis with the BEB score as the criterion, they explained 20% of the variance (Table 13, model 6), almost as much as the unconstrained BFF personality factors (21%). The global TEIQue and ECA scores explained only 5% of the variance. It may be argued that since some of the signs of the beta regression coefficients on the EI subfactors were negative (e.g., self-awareness and group awareness), that this constitutes a violation of EI theory. However, the BFF personality factors also demonstrated similar contradictions with three of the regression coefficients with opposite signs to that predicted by BFF theory applied to EI (McCrae, 2000). Openness and Conscientiousness had negative coefficients while Neuroticism was positively related to the BEB criterion. So, does the sign of the regression coefficient appearing at odds with the theoretical basis of a model render the analysis unacceptable?

Obtaining predictors with an opposite sign to theoretical expectations seems illogical at first glance, but the role that suppressor variables play in fine-tuning regression models is well documented (Howell, 1997). These variables serve to suppress some of the error introduced by the dominant predictors, thus resulting in a better overall regression model fit. For example, Neuroticism is a suppressor variable because it has a negative correlation coefficient with BEB in the zero-order correlation matrix (Table 12), yet it changes to positive in the regression solution. Neuroticism appears to be subtracting some of the error variance from one of the other dominant predictors such as Extraversion. Therefore, there appears to be no theoretical impediment to using the EI test subfactors to determine the incremental variance that EI offers over the BFF personality measures, when predicting the BEB criterion.

By using the unconstrained EI subfactors, an incremental variance of 11% over the BFF personality factors was achieved (model 5, Table 14). This result demonstrates a level of incremental validity where EI offers promising new variables (Mayer et al., 2000b); however, the use of separate EI subfactors does not constitute a single EI construct as proposed by Goleman (1995; Mayer et al., 2001). Mayer et al. (2000b) propose that new variables that provide additional predictions of 1% - 5% over traditional psychometric measures can have meaningful outcomes and

are worthy of further research. The results achieved in this study constitutes evidence for EI theory providing promising new variables for predicting social outcomes.

EI as a predictor of first-semester psychology scores (H3)

At first glance, the global EI scores do not appear to be significant predictors of academic results, as the TEIQue and the ECA have low and insignificant correlations with the PSY104 criterion of 0.18 and -0.03 (Table 12), representing the variance explained of 3% and 0.001% respectively. However, as other predictors are added, EI increases its influence and becomes an important predictor variable for first-year psychology marks (PSY104). While the 30-item TEIQue test alone predicts only 3% of the variance of the PSY104 marks, the 20-item Word reasoning task explains 15% of the variance of the PSY104 marks ($r = .39$). The Word test is used a marker for general intelligence (IQ) in this study.

The Word reasoning score explains a large percentage of the PSY104 score because verbal intelligence has long been associated with academic success. Nevertheless, proponents of EI have suggested that first-year university success is more about adapting to an unstructured environment, managing emotions, making friends and developing independent study habits, than it is about academic intelligence (e.g., Schutte, 1998; Brackett & Mayer, 2003; Petrides, Frederickson & Furnham, 2004). To set a baseline to determine whether EI has any incremental predictive power, the Word score plus the BFF personality factors were entered into a regression analysis which explained 23% of the variance (Table 15, model 2).

Adding the EI subfactors to the baseline as predictors (Table 15, model 6) resulted in a significant 11% of incremental variance explained ($p = 0.04$). This final model (6), with an overall variance explained of 34%, is a useful model in terms of predicting academic performance using the current EI questionnaires.

At this point it is worthwhile to consider the contributions of some of the other tests. Adding Motivation to the predictors (Table 15, model 7) increases the variance explained in PSY104 by another 11%, up to a total of 45%. This simple 15-item motivation scale is a valuable addition to the predictors as it explains as much variance as the 84-item EI subscales. It is of interest to note that Motivation has a negative beta weight in the regression analysis implying that less motivated participants score better in PSY104, with all other factors controlled for. While this seems counterintuitive, a glance at the items under motivation in the Method section shows that the motivation questions address a participant's interest in running a company, making lots of money, and "telling people what to do," which in fact, are probably diametrically opposed to motivators for psychology students who are interested in entering a helping-profession. This suggests that Motivation is not a unitary construct, and there are probably different types of motivation for

different vocations. This serves as a warning that one-size-fits-all EI tests (e.g., Mayer et al., 2002c, Bar-On, 1997; Bradberry, 2003) may not be applicable to all application areas.

In terms of personality, Agreeableness also exhibited a negative beta weight (Table 15) implying that higher Agreeableness scores are associated with lower PSY104 scores. Intuitively, one would assume that the ability to get on well with fellow students would increase success in first-year psychology; however agreeableness such as sharing notes and helping others with their problems does not appear to associate positively with PSY104 results. Perhaps the more successful students who get left doing the majority of the work in group projects change their style so as to be not so agreeable with their social-loafing partners in future. Extraversion is also negatively correlated in a similar manner with PSY104 scores. Perhaps the less extraverted students perform better in PSY104 due to fewer distractions from friends and social activities, and being less agreeable probably leads to fewer party invitations and having fewer friends anyway.

Overall, Mayer (2000b) proposes that new constructs such as EI that provide additional predictions of 1% - 5% over traditional psychometric measures should be considered worthy of further research. In a paper comparing EI tests, Brackett and Mayer (2003) assert that 4% to 9% of additional variance explained can have consequential outcomes. In this regard, the incremental variance of 11% added by the EI subfactors, or 23% total incremental variance if Motivation is included as part of EI suggests that EI has a significant role to play in predicting success in an academic environment.

Replicating the Ford and Tisak study

This study effectively replicated the Ford and Tisak (1983) research to the extent that it found an emotional intelligence (EI) factor with divergent validity from the general intelligence factor. The problem not addressed by Ford and Tisak was whether there was some other, unknown variable outside the scope of the analysis that could be influencing the results obtained. This could have been determined using regression analysis and controlling for personality as I did. This extra factor was discussed briefly in the results section above, where Extraversion turned out to be the pre-eminent predictor of the EI behavioural criterion used here, the BEB interview score. From a regression analysis I conducted on the correlation matrix contained in the Ford & Tisak paper, Hogan's Empathy scale turned out to be the best predictor for their Interview score. However, Hogan's scale has a correlation of 0.63 with MBTI Extraversion, and can be considered as a proxy for Extraversion ($r(42) = 0.63$, $p < 0.000$). It is assumed that Ford and Tisak did not control for BFF personality as it only became a common model in the early 90s (Costa & McCrae, 1985, 1992). Today, it is common practice in EI studies to measure the variance explained by EI as a predictor, after controlling for personality factors (e.g., Davies et al., 1998). Ford and Tisak could have utilised the existing personality scales of the time, for example; the Maudsley Personality Inventory

(Eysenck, 1959), the Eysenck Personality Questionnaire EPQ (Eysenck, 1985), the earlier versions of the 16-PF (Cattell, 1994), or the MBTI (Myers, 1980) to control for the Extraversion factor had they considered it relevant.

Limitations and Strengths of this Study

Overall, the relationship between the brief evaluation of behaviour (BEB) representing a real-world expression of EI, and the EI tests, was not as strong as the Ford and Tisak study (1983) implied that it would be. On the other hand, the ability of the EI subfactors to explain significant incremental variance over general intelligence and personality in predicting first-semester psychology results was a positive result.

The use of non-commercial EI, personality, and other instruments may lead to the criticism that the tests used did not have suitable psychometric properties. While careful to ensure that all scales had acceptable psychometric properties, the cost and the integration of tests into a single battery were two reasons I avoided the commercial offerings such as the MSCEIT and the EQ-i. Another limitation of the study was the potentially limited range of IQ and life experience of the university undergraduate sample, with a mean age of 22.5 years.

On the other hand, the test battery as it stands did enable the production of a clear exposition on the limitations and contradictions when comparing unconstrained BFF personality measures with constrained EI measures which would equally apply to commercial tests. The explanation as to why unconstrained personality factors are dominant when it comes to comparisons with EI is a significant contribution to the EI literature. The high interrater reliability achieved in rating behavioural characteristics was as high as most studies achieve, enabling this criterion-based approach to EI, for which there are few peer-reviewed studies, to be confidently put forward.

Overall, the Ford and Tisak (1983) study was updated to reflect the current paradigm of Emotional Intelligence instead of Social Intelligence, and insights into the limitations of that study were raised. The lower-level EI subfactors were shown to have more predictive power than the global EI scores raising doubt about the assertion that there is a unified EI factor (c.f., Goleman, 1995; Mayer et al., 1999, 2001, 2002c; Bar-On, 1997, 2000). Overall, the EI subfactors were able to explain significant incremental variance over the established constructs of personality and intelligence when predicting a behavioural outcome, and academic results. EI in this study has thus demonstrated significant incremental variance over general intelligence and personality in predicting mature behaviour and academic performance. This evidence provides motivation to explore the EI construct further. EI appears to be a valuable new construct that is worthy of further research and integration into mainstream psychological constructs.

Chapter 4

Study 2 (Chapter 3) revealed that the TEIQue global score, or the ECA global score separately, could explain 5% or 2% respectively of the variance in the behavioural criterion of social maturity obtained from the brief evaluation of behaviour (BEB) - the videotaped interview. But after controlling for personality, the TEIQue only explained 1% of incremental variance. On the other hand, the five factors of personality were able to explain 21% of the variance indicating that the global EI scores of both the TEIQue and ECA were inferior predictors to the well-established big five factor (BFF) personality construct. This finding is congruent with previous research where self-report (trait) EI has been found to be closely related to the personality domain (e.g., Davies, Stankov & Roberts, 1998; Petrides & Furnham, 2000a, 2001, 2004; Matthews, Zeidner & Roberts, 2002).

However, when the EI subfactors were treated as an independent, loosely-related family of constructs (e.g., Ciarrochi & Godsell, 2006; Petrides & Furnham, 2000a; Matthews, Zeidner & Roberts, 2004), and were entered as independent predictors into a regression analysis, the EI subfactors explained a significant 11% of variance in the BEB criterion after controlling for personality. The results from Study 2 therefore indicate that EI, as operationalised by the TEIQue and ECA, is not a unified intelligence, and thus cannot be represented adequately by a total (global) EI score. However, the finding that EI subfactors can explain significant incremental variance over personality scores indicates that EI subfactors may be usefully employed to maximise the variance explained in a criterion. Using the subfactors in EI tests is similar to using other multi-factorial inventories such as the 16-PF or the MMPI to predict outcomes.

Nonetheless, questions arise as to whether the findings in Study 2 were due to the particular EI tests employed, or the criterion employed (BEB). Some may argue that it was an anomaly produced by a student population with an above-average but narrow IQ range, and mean age of 22.6 years indicating a potential lack of worldly experience. It is with these possible limitations in mind that Study 3 was embarked upon using the well-recognised EQ-i test, a sample of experienced employees with a mean age of 37-years, and a real-world criterion of sales-performance.

Emotional Competence, based on abilities from the EI domain of skills is the key to "outstanding performance at work" (Goleman, 2001b, p. 27). Emotional competencies are learned skills; however they require a general ability in EI before they can be developed (Goleman, 1998, 2001b). Within a particular job category such as sales, EI will be a stronger predictor of the "star salesperson" than IQ (Goleman, 2001b, p. 24). IQ only explains up to 20% of the performance variation in job success, leaving other factors, specifically EI, to account for the remaining 80% of the variation in performance (Goleman, 1995, 1998). These claims appear compelling to

organizations attempting to improve the productivity and performance of their staff through Emotional Competence training, and better selection procedures for new staff.

When comparing the performance of the average employee with those performers in the top 15% it was found that the top employees were 20% to 120% more productive than an average employee (Hunter, Schmidt & Judiesch, 1990, cited in Goleman, 2001b). Sales vocations fall into the high range, where top performers typically sell more than twice the value of average performers. This can lead to an outstanding return on investment for the company involved where in one example cited, a top performer selling 120% more than average performers returned additional value equivalent to 88 times his current salary (Goleman, 2001b).

This productivity increase was also reported in a study involving 44 companies such as AT&T, IBM, and PepsiCo, where it was found that salesmen in the top performing 10% group sold more than twice that achieved by average performers (Goleman, 1998). The competencies embodied by the top performers were not IQ related, but were found to be soft interpersonal skills such as Initiative, Empathy and Influence. For want of a general definition of these soft skills, they are usually categorised as EI skills, especially by advocates for the EI concept (e.g., Goleman, 1998; Cherniss, 2004; Boyatzis, Goleman & Rhee, 2000; Bar-On, 1997).

The studies cited above only loosely associate the performance of the top sales performers with EI, because EI is broadly defined as "other characteristics" that are not related to IQ (Goleman, 1995, p. 34). Watkin (2000) asserts that EI, not IQ, is the best predictor of superior performance, superior leadership, and the key to emotionally intelligent organizations; yet he did not supply rigorous definitions or empirical results to back up these assertions. Cherniss (2000, 2001) claims that Emotionally Intelligent managers have less employee turnover, develop staff in a superior manner, promote emotionally intelligent teamwork, and enable people to perform more effectively; again without any empirical data supplied to back up these claims. Cherniss (2004) reported on a Hay/McBer study that indicated that a national insurance company found that representatives that were strong in at least 5 of 8 emotional competencies sold policies twice the value of the remaining representatives. This study was conducted in a closed environment, and no statistics or definitions of EI were provided to be able to examine this claim in detail. Another study using a home-grown EI test and 11 sales professionals from diverse industries ranging from pharmaceuticals to uniform rental reported "a possible link between sales performance and EI" but included no statistical data to support this assertion (Deeter-Schmelz & Sojka, 2003, p. 217).

A somewhat more informative study by Roselle et al. found that sales performance of 103 medical-device sales-people had a low ($r = .20$) correlation with EI as measured by the Schutte et al. (1998) AES inventory (Rozelle, Pettijohn & Parker, 2004). This correlation represents a variance of 4% explained by the EI test which is far below the optimistic estimates of up to 80% of the

variance proposed by Goleman (1995, 1998). Nonetheless, the Rozelle et al. study represents a transparent statement of the relationship of EI with sales performance in a peer-reviewed journal containing statistical information from which we can draw our own conclusions.

However, apart from the last article cited, the conclusion drawn from this brief review of the literature on the topic is that an array of claims have been made without a clear operational definition of EI, without peer-review or published statistical data to back up the claims (Matthews et al., 2002, 2004).

The Bandwidth-Fidelity dilemma

The current study provided the opportunity of surveying up to 100 Sales Representatives with a battery of tests, such as EI and miscellaneous other factors that were assumed to be predictive of sales performance. However, because this survey was an imposition on the already busy Sales Representatives, management wished to keep the time required to complete the battery to less than 90-minutes. This presented a problem, since a battery of complete tests such as the MSCEIT and EQ-i, a full BFF personality test such as the NEO-PIR (Costa & McCrae, 1992), and a full intelligence test such as the WAIS-III (Wechsler, 1997) would have taken up to three hours to complete, and could not be delivered fully over the internet. Thus, Bandwidth (a large number of psychological markers) was chosen over fidelity (more reliable, accurate and established psychological tests) (Cronbach & Gleser, 1965). After the predictive properties of each test was established in this initial study, the long-term strategy was to use tests found to be predictive for personnel selection. The sales performance criterion was consolidated by the corporate sales executive from individual staff performance data supplied by the field sales managers.

Choosing tests for this study

The first step in establishing the link between sales performance and EI is to decide on the best way to measure EI. Due to the lack of agreement between EI theorists there are two classes of EI tests currently being offered: the performance or ability model of EI such as the MSCEIT (Mayer, Salovey & Caruso, 2002c), and the trait or self-report model of EI such as the TEIQue (Petrides & Furnham, 2001, 2003). The Bar-On EQ-i self-report is a trait model of EI (Bar-On, 1997, 1997a) that has also been classified as a mixed model because it contains a mix of 15 different constructs pertaining to EI within the one test battery (Matthews et al., 2002, 2004). Unfortunately, the MSCEIT ability model and the EQ-i trait or mixed model of EI do not exhibit convergent validity as they only share 4% of common variance ($r = .21$), thus demonstrating that they are not measuring the same construct of EI (Brackett & Mayer, 2003). While the MSCEIT 8-factor ability model of EI is attractive from the technical point of view that it tests emotional competency with a range of veridically scored tasks, there are some concerns that it does not have

adequate predictive validity (Bradberry & Greaves, 2004). The EQ-i on the other hand, while being a self-report measure, does have 15 factors contributing to the total EQ score, thereby increasing the chance that some of the factors will prove to be predictive of sales performance (e.g., Bar-On, 2006a).

For this study, the EQ-i was chosen as the model for EI due to its wide acceptance and usage, and the TEIQue was also included to investigate the psychometric properties of a short, 30-item trait EI test. While a full ability EI test such as the MSCEIT was not included, some performance EI tasks such as facial and vocal emotion recognition were included in the battery for testing the predictive properties of performance EI factors over trait EI factors.

In addition to the EQ-i and the TEIQue, the Faces and Voices performance EI tests, a range of miscellaneous tests were added to ensure a comprehensive array of psychological constructs were applied to the task of predicting the sales performance ranking criterion. These additional tests were: (i) A Big Five Factor (BFF) personality battery measuring the five factors of: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism (Saucier, 1994); (ii) markers of general intelligence (IQ) consisting of verbal reasoning and numeric reasoning tasks (Stankov & Roberts, 2001); (iii) miscellaneous scales such as Motivation, Optimism, Impulse control, Empathy, Machiavellianism, and Mechanical aptitude. The mechanical aptitude test was requested by management as it was believed that mechanical ability was a key competency for the Sales Representatives marketing bearings and other mechanical components.

The tests were packaged into a single battery delivered through the internet, with a manual step required to sign on to the Multi-Health Systems (MHS) site to complete the EQ-i test. The MHS system did not provide any facility to integrate the EQ-i test into a single test battery and this introduced the problem of loss of data due to participants having to manually enter the MHS website and complete the EQ-i. Unfortunately, this lack of integration resulted in a loss of approximately 10% of data due to incomplete test results being received. The Sales Representatives were not given any positive incentive to complete the battery; however, the importance of the study was impressed upon them by the Sales Director; this being more of a stick than a carrot approach to motivation.

The overall aim of the study was to determine if EI was a major predictor of success in Sales Representatives, and if results were equivocal, to determine what specific EI factors predicted superior performance. In addition, the predictive properties and the contribution of miscellaneous tests such as personality, general intelligence, and the other tests listed above were explored to determine gaps and shortcomings in the comprehensiveness of EI batteries as predictors. An additional and useful offshoot of this study recognised by the management was to provide guidance

into what characteristics to look for in potential new employees. This was envisaged to consist of a subset of the larger test battery that could be administered on-demand to job applicants.

Summary of aims and hypotheses

The overall aim of this study was to establish the adequacy or otherwise of EI as a predictor of job performance in an organizational setting. The study was designed to confirm or refute the claim that EI can explain a large proportion of the 80% of variance remaining after controlling for IQ (Goleman, 1995). The overall aims can be summarised in five formal hypotheses:

H1: Total EI scores will be significant predictors of the sales performance ranking criterion.

H2: Total EI scores will be better predictors of the sales performance ranking criterion than total IQ scores.

H3: Selected EI subfactor scores will be significant predictors of the sales performance ranking criterion.

H4: Selected EI subfactor scores *and* miscellaneous test scores will explain an even greater amount of variance in the sales performance ranking criterion.

H5: Certain tests will identify superior performers.

Method

Seventy one male technical Sales Representatives from a major industrial bearing and power transmission equipment company in Australia took part in this study sponsored by the Sales Director, the CEO, and the Human Resources department. Overall, a total of 110 Sales Representatives were invited to participate, however the final participation rate of 71 staff (65%) was lower than expected due to remote access computer difficulties, some half-completed surveys leading to incomplete data, and in some cases, Sales Representatives refusing to spend the time to do the survey. The Sales Representatives service major industrial and mining companies through offices located around Australia. The Sales Representative participants were requested by the Sales Director to complete this battery of tests, after-hours, on-location on their own office computer. The participants were advised that the individual's results from the study would remain anonymous, and that the overall results would be used to recruit new Sales Representatives who embodied the characteristics found to be related to success in the current population.

Their ages ranged from 23 to 63 years with a mean of 37.03 years (SD = 8.78). Ninety nine percent of the participants stated that English was their first language, with one participants having Macedonian as his first language. The country of birth of 96% of the participants was Australia, the UK or the USA, with three participants from Italy, Portugal, and Shri-Lanka respectively. In

summary, this sample was, with few exceptions, a unified group of Western-born participants, with English as their first language.

Tasks in this Battery of Tests

Two Emotional Intelligence (EI) tests, the EQ-i, and the TEIQue-SF, plus a variety of personality, general intelligence, and miscellaneous tests that were expected to correlate with sales performance were administered. In addition, a ranking of overall individual performance was prepared by the Sales Managers and consolidated by the Sales Director. This ranking was transformed to a normal distribution under SPSS and used as a criterion for evaluating the effectiveness of each predictor variable in the test battery. A summary of the families of tasks in the test battery is shown below in Table 18.

Table 18

Battery of tasks performed by each participant in this study with the ranking of sales performance criterion of the individual as determined by the manager as the final metric.

Task	Type	Author/Source
Emotional Intelligence tests		
1. EQ-i	Mixed EI	Bar-On (1997)
2. TEIQue-SF	Self-report EI	Petrides et al. (2002)
BFF Personality Inventory		
3. Mini-Markers	Adjective checklist	Saucier (1994)
General Intelligence		
4. Word reasoning	Verbal (*gc*)	Stankov & Roberts (2001)
5. Numerical reasoning	Numeric (*gf*)	Stankov & Roberts (2001)
Miscellaneous tests		
6. Faces	Facial emotions	Ekman (2003)
7. Voices	Vocal affect	Nowicki & Duke (2001)
8. Motivation	Self-report	Various sources
9. Optimism	Self-report	Seligman (1990)
10. Impulse control	Self-report	Tellegen MPQ (1982)
11. Empathy	Self-report	Davis (1980)
12. Mach IV	Machiavellianism	Christie & Geis (1970)
13. Mechanical Aptitude	Practical intelligence	Reid (2006)
Criterion		
14. Ranking of sales performance	Management ranking	Sales Managers

Biographical Details

The first questionnaire was an on-line survey that contained all tests except the EQ-i; the participants were required to manually sign on to the MHS site to complete the EQ-i. The first questionnaire collected the country of birth and first-language of the participants. All participants were male.

Emotional Intelligence tests

Bar-On Emotional Intelligence Inventory (EQ-i). The EQ-i (Bar-On, 1997, 1997a) is a 133-item self-report measure consisting of a mixed array of 15 trait questionnaires such as Assertiveness and Optimism that are combined to produce a Total Emotional Quotient score, the Total EQ score. The items consist statements such as: "*I like everyone I meet*" that are rated on a Likert scale where, 1 = "*Very seldom true of me*" and 5 = "*Very often true of me*." The fifteen factorial components are combined to produce five composite scales: Intrapersonal EQ, Interpersonal EQ, Stress management EQ, Adaptability EQ, and General Mood EQ.

Although the EQ-i scoring keys are proprietary and all tests must be scored through MHS, the scoring is more straightforward because it does not use consensus scoring like the MSCEIT. The results can be therefore interpreted further if the researcher wishes to explore them in detail. The actual participant responses for each item are supplied along with the total scores for the EQ-i to enable further analysis if required, although this was not done.

Trait Emotional Intelligence Questionnaire – Short Form (TEIQue-SF). The TEIQue-SF is a 30-item, short-form version of the 144 item TEIQue, Version 1 (Petrides, Perez & Furnham, 2003, 2004). The TEIQue-SF was designed to yield a global EI score and does not provide the 15-subfactors that are available in the 144-item TEIQue. The TEIQue-SF is built on a 4-factor theoretical structure: emotional Well-being, Self-control, Emotionality and Sociability.

The questions such as, "*I usually find it difficult to regulate my emotions*" are responded to on a 7-point Likert scale ranging from: 1 = "*Completely disagree*" to 7 = "*Completely agree*." The scoring key for the global trait EI score is provided by the authors on the questionnaire documentation, and the scoring key for the four factors was obtained from the author K. V. Petrides (personal communication, March 1, 2006). For simplicity, this paper will omit the "-SF" suffix and simply refer to the test as the TEIQue throughout this document.

Facial affect recognition task (Faces). The Micro Emotion Training Tool (METT) was administered as a test of facial affect recognition ability (Ekman, 2003, 2003a). Facial emotion recognition ability is a component of EI (Mayer et al., 2002c). The METT utilises the same stimuli as the Japanese and Caucasian brief affect recognition task (JACBART – Matsumoto et al., 2000). The METT comprises 56 stimuli and presents a brief one-fifteenth of a second emotion, forward

and backward-masked by the same face without the emotional expression. Internal consistency reliability of the test (Cronbach's alpha) is 0.86 for 1/15th of a second exposure (Matsumoto et al., 2000). After each stimulus is presented, the participant is asked to click a button to rate the expression as one of the following seven emotions: *Happiness, Disgust, Sadness, Anger, Surprise, or Fear*. A high score in Faces indicates a superior ability to recognise emotion, and thus contributes to a higher total EI score. Contempt items were removed from the stimuli presented, as a previous study had indicated that there was general confusion between Contempt and Disgust emotions which lowered the overall reliability of the test (Reid, 2004).

Vocal affect recognition task (Voices). The vocal emotion recognition task utilised ten stimuli from the DANVA2 battery of adult male and female voice clips (Nowicki, 1998; Nowicki & Baum, 1998; Nowicki & Duke, 2001). The vocal-emotion items were created by two professional actors who rendered the simple, neutral sentence, "*I am going out of the room now but I'll be back later*," in a variety of emotions. Participants responded by pressing one of four buttons corresponding to the emotion they identified: *Happy, Sad, Angry,* and *Fearful*. The participants were asked to put on headphones for this task so that they would not disturb other participants. Higher scores on the DANVA2-AP are reputed to be related to measures of social adjustment such as lower depression and social anxiety, and higher self esteem (Nowicki & Duke, 2001).

Big-Five Factor Personality Inventory (BFF)

The BFF personality inventory employed in this study was, Mini-Markers: a brief version of Goldberg's unipolar Big-five markers (Saucier, 1994). This 40-item inventory demonstrates high alpha reliabilities on all five factors: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism in an easy to administer test. For simplicity, the abbreviations of O, C, E, A and N will be used throughout this document to describe the five factors above.

The instructions for questionnaire asks the participant, "*How accurately can you describe yourself*?" It then presents 40 personality adjectives such as, "*Bold*," "*Shy*," and "*Warm*," that the participant rates on a 9-point Likert scale ranging from, 1 = "*Extremely inaccurate*" to 9 = "*Extremely accurate*." The scoring key and factorial structure is provided with the documentation to enable the experimenter to score results. In this study, the Likert scale was reduced to 5-points, still using the same endpoints as described above, to enable comparisons with a larger sample collected at the University using computer readable answer sheets that only had five response options. The reduction of the Likert scale to 5-points did not appear to have an adverse effect on alpha reliabilities.

General intelligence markers (IQ)

While the markers for crystallised intelligence (Gc) and fluid intelligence (Gf) did not constitute a comprehensive IQ test, they were useful markers to compare with EI test results to establish divergent and convergent validity. The general intelligence dimension of individual differences will be referred to as IQ for brevity throughout this study, although it is acknowledged that this is not a complete IQ test.

Word Reasoning (Word). The Word reasoning task was derived from 20-items from the Esoteric Analogies test in the Gf/Gc Quickie Test Battery (Stankov, 1997; Stankov & Roberts, 2001). The participant is presented with a sentence with three key words such as: "*Horse is to Animal as Chair is to*:" and four alternate answers to choose from, such as: "*Furniture, Leg, Person, Table.*" The participant's total score out of 20 is simply the number of items that they answer correctly. The Word task is a marker for verbal skills or crystallised intelligence (*gc*) (Carroll, 1993).

Numeric Reasoning (Odd-one-out). The Odd-one-out task consists of a 10-item subset of questions taken from the Gf/Gc Quickie Test Battery (Stankov & Roberts, 2001). The participant was asked to determine which sequence of four numbers is the odd-one-out, for example:

"*7-4-8-9,*" "*6-5-9-2,*" "*5-4-8-3,*" "*9-4-8-5,*" "*6-4-8-2*."

The rules change for each of the ten items in this task, thus it is a measure of ability to use reasoning to see abstract patterns, which is the defining property of fluid intelligence (*gf*) (Carroll, 1993). While higher reliability could be achieved with 20-items, it was decided that the battery should not appear to be an onerous IQ test; thus the numerical reasoning section was kept short.

Miscellaneous inventories

Motivation. The Motivation test items, designed to test motivation to achieve and succeed, were assembled from a variety of public domain sources. Items were reverse scored and counterbalanced to avoid any effects from a tendency to respond in the affirmative. A five-point Likert scale ranging from: 1 = "*Never*" to 5 = "*Always*" was utilised. Examples of motivation items are as follows:

"*I avoid positions of leadership(r)* "

"*My level of motivation to achieve is high*"

"*I seek a quiet life and avoid challenges (r)* "

The responses were reversed where applicable (*r*) and summed to provide a total Motivation score.

Optimism. The Optimism scale utilised was from the Optimism scale in Seligman's book entitled "Learned Optimism" (1990). Each of the 48-items is provided with a binary choice of answers which the participant must select from. The face validity of the items is low and it is

difficult to anticipate the correct answers. The theory is quite complex with 10 factors and subfactors such as: *Permanence Bad, Permanence Good, Pervasiveness Bad, Pervasiveness Good, Hope, Personal Bad, Personal Good, Total Bad, Total Good, and Overall Score* (Good - Bad). An example of an item is show below:

1. The project you are in charge of is a great success.

A. *I kept a close watch over everyone's work.*

B. *Everyone devoted a lot of time and energy to it.*

The correct choice for the above is (A) which adds one point to the *Pervasiveness Good* score. Often the obvious or politically correct answer such as (B) is the incorrect one leading participants to obtain lower than expected overall optimism scores, depending on their orientation to individualism or collectivism. The responses were scored according to the keys provided, and the ten scores were calculated for each participant.

Impulse control. The 16-items that comprise the Impulse control inventory were taken from Tellegen's Multidimensional personality questionnaire (MPQ) (Tellegen, 1982). Items were scored on a 5-point Likert scale where: 1 = "*Never true*" to 5 = "*Always true*." Examples of Impulse control items are as follows:

"*I often act without thinking (r)*"

"*I usually make up my mind through careful reasoning*"

"*I often act on the spur of the moment (r)*"

The responses were reversed where applicable (*r*) and summed to provide a total Impulse control score.

Empathy. The 13 Empathy items were taken from the 27-item Interpersonal Reactivity Index (Davis, 1980, 1983) with items such as, "*I really get involved with the feelings of the characters in a novel*" omitted because they did not apply to corporate activities. Items were scored on a 5-point Likert scale where: 1 = "*Strongly disagree*" to 5 = "*Strongly agree*." Examples of Empathy items are as follows:

"*I don't feel very sorry for other people when they are having problems (r)*"

"*I believe that there are two sides to every question and try to look at them both*"

"*I sometimes find it difficult to see things from the other person's point of view (r)*"

The responses were reversed where applicable (*r*) and summed to provide a total Empathy score.

Machiavellianism (Mach IV) scale. The Mach IV scale (Christie & Geis, 1970; Panitz, 1989) was included in the battery to determine if Machiavellianism played a role in successful sales performance. The philosophy of Niccolo Machiavelli (1469-1527) espoused the concept that success did not depend on virtue and honour, rather it depended on manipulation, such as telling the people what they wanted to hear, regardless of the truth. Thus, the relationship between

Machiavellian attitudes and successful sales performance were explored with this short, 20-item scale. Items were scored on a 5-point Likert scale where: 1 = "*Strongly disagree*" to 5 = "*Strongly agree*." Examples of scale items are as follows:

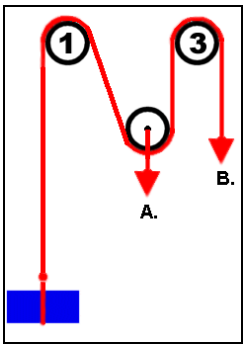
"*Never tell anyone the real reason you did something unless it is useful to you*"

"*The best way to handle people is to tell them what they want to hear*"

"*One should take action only when it is morally right (r)*"

The responses were reversed where applicable (*r*) and summed to provide a total Empathy score.

Mechanical Aptitude test. This short, 10-item mechanical aptitude scale was assembled from a variety of public domain diagrams available on the internet and other sources. For each item, a mechanical diagram was presented and the participant given four options to choose from. An example of an item is shown below:



"What option below would require the least effort to lift the weight?"

A. *Attach A. to a fixed point and pull on B.*

B. *Attach B. to a fixed point and pull on A.*

C. *Pull on A. and B. simultaneously*

D. *All options require the same effort*

The correct responses were summed to provide a total Mechanical Aptitude score.

Ranking of performance of Sales Representatives

The Sales Managers in all states of Australia were asked to produce a ranked-list of their Sales Representatives, ordered from the highest performer to the lowest performer. This ranking was not based on dollar volume of sales due to the disparate income from different regions, but rather, on consistent performance with respect to the opportunity in the sales territory. These reports were consolidated into one list by the Sales Director. The ranking was transformed into normal distribution using the *transform to normal scores* function of SPSS to enable normal/parametric

statistics to be calculated. Thus, the highest performer was given a score of +2.37, while the lowest performer was given a transformed normal ranking score of -2.37.

In addition to the sales performance ranking, the individual managers also allocated a Performance Rating Scale (PRS) number (1 – 4) for each participant, where: 1 = *Superior* performer, 2 = *Average* performer, 3 = *Below-average* performer, and, 4 = *Substandard* performer. Total numbers in each group was 14, 34, 14 and 8 respectively.

Test Procedure

Participants took the tests after-hours, at their own desks in the work environment. This reduced interruptions from telephone calls and business-hours activities. The battery of tests consisting of the two on-line questionnaires typically took 60 to 90-minutes to complete.

Results

Descriptive Statistics

Table 19

Descriptive statistics for the battery of tests (N=71), with coefficient alpha, and standardised data from the original test-developers where available.

	This Study			Standardised Data		
Test	Mean	SD	Alpha	Mean	SD	Alpha
EQ-i Total EQ	96.71	14.01	(1)	100	15	.80
TEIQUE Total EI	5.10	.81	.86	5.07	.65	.86
Openness	28.41	4.47	.72	30.9	4.6	.78
Conscientiousness	31.83	4.56	.83	28.4	5.1	.78
Extraversion	28.27	4.43	.68	26.9	5.4	.81
Agreeableness	31.10	4.03	.69	32.6	4.4	.81
Neuroticism	16.94	4.47	.78	20.4	4.8	.79
Word reasoning score (*gc*)	.71	.16	.72	0.80	.11	.71
Numeric reasoning score (*gf*)	.75	.20	.68	0.73	.27	.87
IQ average	.73	.16	.80	(2)		
Faces	.72	.10	.67	(2)		.86
Voices	.62	.21	.67	(2)		.78
Motivation	3.65	.47	.66	3.36	.48	.69
Optimism	2.68	4.40	.64	(2)		
Impulse control	3.73	.50	.80	3.33	.57	.84
Empathy	3.59	.55	.66	3.76	.61	.77
Mach IV scale	2.59	.54	.79	(2)		
Mechanical aptitude	6.90	1.80	.82	(2)		
Ranking of performance	0.00	1.0	(3)	(2)		

Notes: (1) Alpha reliabilities are not provided by MHS with the scored sample. (2) This is a new scale and there is no standardised data available. (3) The performance ranking score is a single standardised (*z-score*) rating and thus alpha reliabilities cannot be calculated.

All skew and kurtosis statistics were within the range of plus one to minus one; the maximum statistic being 0.83 indicating that the data did not require transformation before further analysis (Dekker, 2005). Some scales such as Faces, Voices, Motivation and Optimism exhibited low alpha reliabilities due to the limited number of items in each scale; this was the result of trading fidelity off for bandwidth.

Correlation Matrix

Table 20

Relationships between major variables in the test battery.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. EQ-i Total EQ	1																		
2. TEIQ total EI	.44**	1																	
3. Openness	.22	.19	1																
4. Conscientiousness	.44**	.12	.15	1															
5. Extraversion	.41**	.21	.30*	.28*	1														
6. Agreeableness	.32**	.14	.20	.22	-.04	1													
7. Neuroticism	-.52**	-.40**	.11	-.31**	-.27*	-.32**	1												
8. Word reasoning	.09	.06	.12	.04	.07	.08	-.21	1											
9. Numeric reasoning	.25*	.04	-.11	.26*	.12	.14	-.37**	.49**	1										
10. IQ score	.21	.06	-.01	.19	.11	.13	-.35**	.82**	.90**	1									
11. Faces	.03	.03	-.07	.04	.06	-.06	-.23	.43**	.39**	.47**	1								
12. Voices	.08	.10	-.08	-.02	-.16	.30*	-.09	-.03	.01	-.01	.12	1							
13. Motivation	.39**	.27*	.39**	.27*	.52**	.09	-.23	-.07	.01	-.03	.03	-.05	1						
14. Optimism	.18	.26*	.18	.30*	.20	.25*	-.10	-.05	-.15	-.12	-.31*	.00	.22	1					
15. Impulse control	.54**	.27*	.05	.56**	.33**	.24*	-.43**	.05	.19	.15	-.04	-.16	.23	.25*	1				
16. Empathy	.56**	.39**	.22	.10	.15	.53**	-.40**	.06	.09	.10	-.07	.13	.30*	.17	.35**	1			
17. Mach IV	-.32**	-.33**	-.03	-.20	-.16	-.22	.32**	-.16	-.19	-.21	.13	.03	-.05	-.12	-.23	-.32**	1		
18. Mechanical Aptitude	.24	.25	-.14	.08	.04	.10	-.07	.21	.28*	.30*	.22	.23	.03	-.08	.19	.18	-.07	1	
19. Age	.17	.04	.06	-.01	.02	.12	.04	-.10	-.17	-.16	-.30*	.15	-.04	.04	.02	.11	.02	-.05	1
20. Performance ranking	.10	.06	.10	-.04	.36**	-.08	-.04	.11	.00	.06	-.07	.08	.22	.05	-.15	.05	-.19	.11	.34**

Note: ** indicates that correlation is significant at the 0.01 level (2-tailed), * indicates significance at the 0.05 level (2-tailed).

The correlation matrix (Table 20) indicates that only two scores, Extraversion and Age, correlate significantly ($p < .01$) with the ranking of performance criterion. Neither the EQ-i, nor the TEIQue global scores correlate significantly with the sales performance ranking criterion, and both exhibit small (0.10 - 0.29) or trivial (<.09) correlations of 0.10, and 0.06 respectively (Matthews, et al., 2004). The issue of predictive validity of the EI tests will be explored in more detail in the regression analysis.

Relationship between EQ-i subfactors and the criterion

Table 21

EQ-i and TEIQue subfactor correlations with the sales performance ranking criterion

EI Subfactor	Sales Performance Ranking Correlation
EQ-i	
Self regard	.18
Emotional SA	-.001
Assertiveness	.18
Independence	-.12
Self actualization	.12
Empathy	.22
Social responsibility	.16
Interpersonal relations	.02
Stress tolerance	.15
Impulse control	-.12
Reality testing	.002
Flexibility	-.02
Problem solving	.03
Optimism	.28*
Happiness	.20
TEIQue	
Well-being	.05
Self-control	-.09
Emotionality	-.02
Sociability	.27*

* = Correlation is significant at the 0.05 level (2-tailed).

Table 21 shows that not all EQ-i or TEIQue subfactors are positively correlated with excellent performance as represented by the sales performance ranking criterion. The EQ-i factors of Independence, Impulse control, and Flexibility are negatively related with the sales performance ranking criterion indicating that high performers obtain lower scores on these factors. This is also

true for the TEIQue subfactors of Self-control and Emotionality which have negative correlation coefficients. The only significant correlation with the criterion in Table 21 is the EQ-i factor of Optimism and the TEIQue factor of Sociability.

H1: Total EI scores will predict the Performance Ranking

Hypothesis one predicted that total EI scores from the EQ-i and the TEIQue would be significant predictors of the sales performance ranking criterion. The correlation matrix shown in Table 20 indicates that the EQ-i total EQ score and the TEIQue total EI score correlations of 0.10 and 0.06 respectively were insignificant and trivial correlations (Matthews et al., 2004) with the sales performance ranking criterion, representing a variance explained of 1% and 0.36% respectively.

Therefore, hypothesis one is rejected; total EI scores did not predict the sales performance ranking criterion.

H2: Total EI scores will be better predictors than IQ scores

Hypothesis two predicted that total EI scores would be better predictors of performance than IQ scores. The IQ score, made up of Word reasoning and Numeric reasoning task scores, exhibited insignificant and trivial correlations of 0.11 and 0.00 respectively with the sales performance ranking criterion (Table 20). Word reasoning, or Verbal reasoning, appears the most important general intelligence skill because it has the highest correlation (0.11), however; these correlations are insignificant and no concrete conclusions can be drawn. The total EI scores also exhibited insignificant and trivial correlations of 0.10 and 0.06 respectively with the performance ranking criterion.

Thus, hypothesis two is rejected on the basis that all correlations were insignificant and trivial, and the hypothesis could not be demonstrated to be true.

H3: Selected EI subfactor scores will be significant predictors of performance

Hypothesis three proposed that selected EI subfactor scores would be significant predictors of the performance ranking criterion. To select the significant predictors, all EI subfactors were entered into a stepwise regression analysis using the "backward" predictor removal method to obtain the best model possible. The predictors remaining from this stepwise process were entered into a regression analysis using the "enter" method to confirm the model was indeed significant. This model appears as Model 3 in Table 22.

Table 22 summarises the variance explained by various models: (1) the Total EQ-i score, (2) the EQ-i subscale of General Mood, (3) selected EI subfactors that explain 38%, and, (4) the addition of miscellaneous variables that bring the variance explained up to 52% of the sales performance ranking criterion.

Table 22

The variance explained by selected EI and miscellaneous factors.

Model	Predictors for the Sales Performance Ranking Criterion	Beta Estimates	Parameter p – value	Overall Model Fit p – value	Total variance explained R^2
1.	Total EQ-i score	.10	.44	.44	1%
2.	General mood EQ-i subscale	.27	.03	.03	7%
3.	1. EQ-i Independence	-.31	.02	<.0001	38%
	2. EQ-i Interpersonal relationship	-.40	.01		
	3. EQ-i Stress tolerance	.43	.004		
	4. EQ-i Happiness	.35	.03		
	5. TEIQue Self-control	-.54	.001		
	6. TEIQue Sociability	.64	.001		
4.	1. EQ-i Independence	-.33	.007	<.001	52%
	2. EQ-i Interpersonal relationship	-.38	.008		
	3. EQ-i Stress tolerance	.32	.02		
	4. EQ-i Happiness	.30	.04		
	5. TEIQue Self-control	-.41	.005		
	6. TEIQue Sociability	.47	.001		
	7. Age	.34	.001		
	8. BFF Extraversion	.24	.03		

Note that all models except 1 are significant at $p < .05$. Also note that some of the EI subfactors such as EQ-i Independence, EQ-i Interpersonal relationship, and TEIQue Self-control have negative relationships with the performance criterion. This indicates, for example, that less Independence results in higher performance, when all other subfactors are held constant.

Model 3 is superior in terms of predictive properties over the total EQ-i score (model 1) that explained a statistically insignificant 1% of variance, and the EQ-i subscales, of which only General Mood was significant at $p < .05$, explaining 7% of variance (model 2). The TEIQue total EI only explained 0.36% ($r = .06$) of the variance in the sales performance ranking criterion making it the least-predictive total EQ score. Thus, the EI subfactor scores are far more predictive of performance than the total or global EI scores for either test. It is noteworthy that Independence, Interpersonal relationship, and Self-control are all negatively related to the success criterion, sales performance ranking. This will be explored further in the discussion.

Hypothesis three is therefore supported; selected EI subfactor scores were significant predictors of the sales performance ranking criterion.

H4: Selected EI and miscellaneous variables will be significant predictors

Hypothesis four proposed that selected EI subfactor scores *and* miscellaneous variables would explain an even greater amount of variance in the performance ranking criterion. The rationale behind this hypothesis was to examine whether EI tests explain all the variance in the criterion, or whether there was significant variance left over that could be explained by other non-EI tests and variables.

To select the significant predictors, the EQ-i subfactor predictors from Table 22, model 3, were entered into a hierarchical regression analysis (step 1) and all additional non-EI tests were added with a "stepwise" predictor addition method (step 2). The predictors obtained from this stepwise process were then entered into a regression analysis using a 2-step "enter" method to confirm the model and the predictors were all significant.

The results shown in Table 22, model 4, above indicate that the traditional predictors of Age and Extraversion explain significant incremental variance beyond both TEIQue and EQ-i subfactors. This demonstrates that there are other miscellaneous variables that can explain significant incremental variance beyond the EI tests utilised in this study, and thus hypothesis four is accepted.

H5: Tests that predict the difference between Average and Excellent performers

Hypothesis five postulated that there would be significant differences in certain tests results between Superior and Substandard performers on the Performance Rating Scale (PRS) criterion. To determine which tests were predictive, an independent samples *t*-test was conducted on all 34 tests and EI subfactors utilised in the battery. Only those tests that demonstrated significant differences ($p < .05$) between *Superior* (1) and *Substandard* (4) performers are shown below.

Table 23

Tests that discriminate significantly between Superior (1) and Substandard (4) performers as defined by the performance rating scale (PRS) criterion.

Variable	*t*	*df*	*p* - value (2-tailed)	Mean Difference	Mean score PRS = 1	Mean score PRS = 4
Age	4.49	20	.0002	11.52	44.64	33.13
Sociability	3.03	20	.008	0.97	5.29	4.31
Extraversion	3.82	30	.001	5.18	30.43	25.25

Note that all differences are statistically significant at $p < .01$, and are all in the theoretically correct direction; with *Superior* performers (1) scoring higher on all three measures. Clearly, Age

(p = .0002) is an asset in this job role, as is Extraversion and Sociability. Below is a graphical representation of the data in terms of average scores obtained by performance groups 1 to 4.

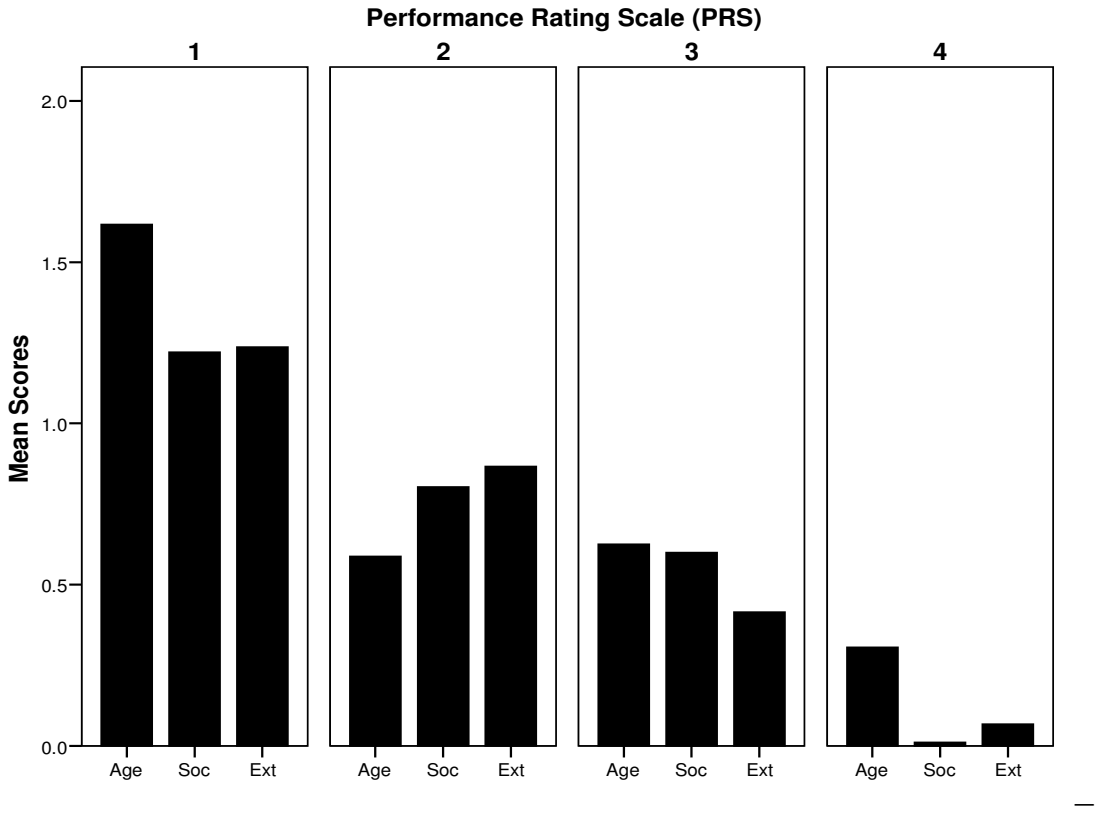


Figure 7

Figure 7 demonstrates how mean scores of Age, Sociability (Soc), and Extraversion (Ext) diminish as the Performance Rating Scale (PRS) goes from Superior (1) to Substandard (4) performance.

Discussion

The five hypotheses for this study were: (i) that total EI scores will be significant predictors of the sales performance ranking criterion; (ii) total EI scores will be better predictors of the sales performance ranking criterion than total IQ scores; (iii) selected EI subfactor scores will be significant predictors of the sales performance ranking criterion; (iv) selected EI subfactor scores *and* miscellaneous test scores will explain an even greater amount of variance in the sales performance ranking criterion; and, (v) certain tests will identify superior performers.

The following discussion will focus on the implications and findings in the light of these hypotheses.

Total EI scores as predictors of performance (H1)

Hypothesis one predicted that total EI scores from the EQ-i and the TEIQue would be significant predictors of the sales performance ranking criterion. The correlation matrix in Table 20

shows that neither the EQ-i or the TEIQue total EI scores were significant predictors of the sales performance ranking criterion. Thus, hypothesis one is rejected; total EI scores were *not* significant predictors of the sales performance ranking criterion.

The variance explained by the EQ-i and the TEIQue, 1% and 0.36% respectively, was disappointing in terms of the promise regarding EI; that it would potentially explain up to 80% of the variance proposed by Goleman (1995, 1998, 2001, 2002). If the study had utilised only one EI test, the efficacy of that particular test would have been in question, however, since two EI tests produced the same trivial results, it appears conclusive that the total EI scores were not predictive of the sales performance ranking criterion. Mayer has set a benchmark that new constructs that explain an additional 5% of variance ($r = .22$) are worthy of further investigation (Mayer, Salovey & Caruso, 2000b). The result obtained in this study falls well short of the Mayer benchmark, let alone the promise of up to 80% proffered by proponents of EI (Goleman, 1995, 1998, 2001a, 2001b).

Perhaps it could be argued that an ability or performance EI test like the MSCEIT may have explained more variance than the self-report tests did. Turning to the performance EI tasks of Facial and Vocal emotion recognition that were included as markers of performance or ability EI, correlations of -.07 and .08 respectively were observed (Table 20). Both of these tasks explain less than 1% of variance each, suggesting that performance tasks, such as Faces in the MSCEIT, would have fared no better than the Faces and Voices performance markers given in this battery. While the Faces correlation of -.07 with the criterion is insignificant, it is also in the wrong direction, indicating that ability to recognise emotions in faces is negatively related to performance. This questions the entire concept of EI where emotional abilities are alleged to be positively related to performance, especially in roles that involve interpersonal relationships such as Sales Representatives (c.f., Spencer, 2001).

Thus, while hypothesis one is firmly rejected, additional questions are raised about the trivial amount of variance explained by total EI scores compared with the promise of 80%, and the issue of performance or ability EI tests actually being negatively related to the sales performance ranking criterion. Overall, total EI scores have not lived up to the promise of EI to explain a significant proportion of the 80% of the variance in job performance (Goleman, 1995, 1998, 2001a).

Total EI scores will be better predictors than IQ scores (H2)

Hypothesis two predicted that total EI scores would be better predictors of the sales performance ranking criterion than total IQ scores. As discussed in the results section, the total EI scores and IQ scores demonstrated trivial and insignificant correlations (Table 20) with the sales ranking performance criterion. Hypothesis two was rejected on the basis that EI scores showed no greater predictive properties than IQ scores.

In terms of actual correlation coefficients, the Verbal reasoning task had a marginally greater correlation with the sales ranking criterion than any of the other EI or IQ scores. This indicates that Verbal intelligence may proffer some advantage for a sales representative, while Numerical reasoning with a correlation of zero appears to offer no such advantage. Overall, the lower than expected and trivial correlations of EI and IQ with the criterion make the comparison between EI and IQ meaningless. Therefore, in a situation such as this where there is no data to decide which the better predictor is, hypothesis two must be rejected due to lack of conclusive results.

Selected EI subfactor scores as significant predictors of performance (H3)

Hypothesis three proposed that selected EI subfactor scores would be significant predictors of the sales performance ranking criterion. To choose the significant EI subfactors a stepwise regression was conducted as described in the results section. This regression analysis yielded six EI factors that explained 38% of the variance in the sales performance ranking criterion. The regression analysis was highly significant ($p < .0001$) leading to an unequivocal acceptance of hypothesis three; that EI subfactors are significant predictors of the sales performance ranking criterion.

In terms of absolute variance explained, the 38% exceeds the 5% criterion proposed by Mayer by a generous margin (Mayer et al., 2000b), and begins to approach the promise of EI to explain up to 80% of the variance available for factors other than IQ (Goleman, 1995).

However, the results are more complex than they first appear. The unearthing of significant EI subfactors using regression has not generally been the focus of EI studies in the past, although Bar-On provides one such study (Bar-On, Handley & Fund, 2006). The theoretical concept that all EI factors are unidirectionally predictive as proposed by the theorists (e.g., Mayer et al., 2002c; Bar-on, 1997; Goleman, 1998, 2001, 2002) has not been discussed in detail before. Most EI studies are content to follow the theoretical assumption that EI subfactors can simply be summed to produce a total EI score. Subfactors are assumed to be additive, and their directionality has been designated on an *a priori* basis. This study provides the data to test empirically whether these assumptions are correct.

Referring back to the six EI subfactors that explained 38% of the variance (Table 22, model 3) it can be seen that three of the subfactors have significant negative beta weights indicating that with all other factors held constant; an increased score in Independence, Interpersonal relationship or Self-control is inversely related to high performance. This pattern of negative directionality is also present in the zero-order correlation matrix (Table 21) with the exception of the EQ-i Interpersonal relationship factor. This finding is also supported in a study where Impulsivity, the opposite pole of Impulse-control and Self-control, was found to be best overall predictor for sales

performance and goal achievement (Jackson, 2001). The concept of superior performance being inversely related to certain EI subfactors has not been explored in detail before in the EI research literature. The appearance of negatively correlated EI subfactors brings into question the assertion that EI is a unitary intelligence directly analogous to general intelligence (e.g., Mayer et al., 1993, 1999; Goleman, 1995).

There are four factors out of 15 (27%) in the EQ-i, or six out of 19 (32%) if the TEIQue factors are included, that have negative correlations with the sales performance ranking criterion (Table 21). Clearly, not all EI subfactors are positively related to superior performance as the EI theorists assume (e.g., Goleman, 1995, 2001; Bar-On, 1997; Mayer et al., 2002c). The mixture of positively and negatively related subfactors when summed to produce a total EI score reduces the ability of the total EI score to explain significant variance. This is clearly demonstrated in this study where the total EQ score of both the EQ-i and the TEIQue only explained 1% or less of the variance.

It may be argued by some that the sales performance ranking criterion used in this study was somehow unsuitable for testing the properties of EI because it was a manager's ranking and subjective in nature. However, in terms of scientific investigation, a theory that proposes that all EI subfactors are positively related to high performance, can be summed to produce a total EI score, and, that "EI meets traditional standards for an intelligence" (Mayer, Caruso, & Salovey, 1999) must be prepared to subject itself to scientific scrutiny and falsification (Popper, 1963). In this example, the proposal that all EI subfactors can be summed because they all relate positively to excellence in the job-performance criterion is falsified on the basis of evidence from the correlation table (Table 21) and the regression analysis (Table 22).

This minor diversion into the theoretical status of EI and falsification does not, however, change hypothesis three that states: selected EI subfactor scores will be significant predictors of the sales performance ranking criterion. It broadens the discussion by showing that some of the significant EI subfactor predictors are negatively related to the criterion of success. For example, the correlation matrix in Table 20 shows that Impulse control is negatively related to the sales performance ranking criterion; Table 21 also indicates that the EQ-i Impulse control and the TEIQue Self control factors are negatively related to sales performance; and, Table 22, models 3 and 4 also show a negative relationship between Self control and the sales performance criterion. This is robust evidence that Impulsivity, the opposite of Impulse control, is valuable characteristic for sales representatives; this finding being confirmed in at least one other study (Jackson, 2001). Thus, hypothesis three is accepted, with the additional knowledge that there are some significant EI subfactors that correlate negatively with the criterion.

Miscellaneous test scores will add incremental variance (H4)

Hypothesis four proposed that selected EI subfactor scores *and* miscellaneous test scores will explain an even greater amount of variance in the sales performance ranking criterion. Table 22, model 4, shows that the addition of Age and the BFF factor Extraversion increases the variance explained in the sales performance ranking criterion from 38% to 52%. This selected set of predictors explaining 52% sustains the promise that "other characteristics" not related to IQ will potentially explain up to 80% of the variance in job performance (Goleman, 1995, p. 34). Age and Extraversion are also significant predictors at the $p < .01$ level on the correlation matrix in Table 20. However, some caveats must apply; Age and Extraversion, while not being IQ constructs, are not available for reclassifying as EI constructs as they have already been defined as demographic and personality variables respectively. However the point is accepted, that *other characteristics* beyond IQ can add significantly to the prediction of the job performance criterion.

The use of Age as a predictor must be approached with caution due to legal concerns regarding privacy and discrimination. Age in this study may be acting as a proxy for years of experience in this industry. It could also represent maturity, stability and focus on the job, where the advancing years also brings into focus the need to work hard to build up a good retirement fund. Either way, the Age variable requires further investigation and decomposition into its relevant subfactors using additional questions such as: years in this industry; years as a sales representative; stability as measured by years in the same location; social and community activities; number of working years ahead before retirement; and enjoyment derived from the job and the company. This is by no means a complete list, but it does provide some direction as to how to proceed to obtain an insight into the subfactors comprising the Age variable.

A high score on the Extraversion predictor indicates an outwardly-directed personality, that enjoys social situations, and who specifically rated themselves highly on items such as: Bold, Energetic, Extraverted and Talkative on the personality questionnaire. They rated self-descriptive terms such as: Shy, Quiet, Bashful and Withdrawn as low on this scale. It is intuitively clear that Extraversion is a useful personality trait for a sales representative who needs to make cold-calls to potential customers, and to be outwardly directed towards often-difficult customers on a daily basis. However, as useful as Extraversion is as a predictor it cannot be subsumed under the EI banner, as it clearly belongs to the BFF personality constructs (Costa & McCrae, 1985, 1992; McCrae & John, 1992).

In summary, there are other factors beyond EI that explain significant incremental variance in the sales performance ranking criterion; thus hypothesis four is unequivocally accepted.

Certain tests will identify superior performers (H5)

This final hypothesis takes a different approach to the research question by looking at the top 15 performers (21%) and comparing their results with the bottom 8 performers (11%). Using this approach it was found that an independent-samples *t*-test identified three predictors as highly significant ($p < .01$): Age, Extraversion and the TEIQue Sociability EI factor. This confirms the robustness of Age and Extraversion as predictors that have already been discussed above.

A high score on the TEIQue Sociability factor indicates a tendency to endorse the high end of the Likert scale (*Completely Agree*) for statements such as: "I can deal effectively with people; I am usually able to influence the way other people feel; and, I would describe myself as a good negotiator." Although the TEIQue scale is labelled Sociability, it seems to be addressing negotiation skills and an ability to influence other people's behaviours. Clearly, this trait would be helpful in achieving positive outcomes during sales contract negotiations. The TEIQue Sociability scale appears to be a robust indicator of performance as it appears consistently through the regression analyses and this current examination of differences between the top and bottom performers.

Hypothesis five is therefore unequivocally accepted; Age, and tests of Sociability and Extraversion identify excellent sales performers.

Limitations of this Study

The first limitation of this study was that the final sample size of 71 participants out of a possible 110 Sales Representatives was smaller than expected by the management. This stresses the need for a positive incentive to do the survey rather than an edict from the Sales Director. In the following study, the incentive to do the survey was paid overtime after the day's work had been completed. Unfortunately, this would not have been applicable to the higher-level job role of the sales representative who manages his own time to achieve his own personal goals (note, there were no female sales representatives in the sample).

Another limitation was not including a performance measure of EI such as the MSCEIT in the battery. Unfortunately, the current battery took up to 90 minutes to complete, and the addition of an extra 30-minutes to do the MSCEIT would have made the battery too long. As discussed, two performance EI tasks, Faces and Voices, were included to determine the predictive properties of ability EI tests.

The ranking of performance of Sales Representatives by the Sales Director may be criticised as being too subjective and not based on hard measures such as dollar volume income. The prime predictors of success that turned out to be sociability and extraversion suggest that likeable personalities may have been ranked higher. Nevertheless, assurance was given that the original ranking was provided by the Field Managers, and simply merged by the Sales Director.

An alternate regression model also explaining 52% of the variance included Optimism as a predictor, suggesting that other potential regression models may be possible. This is not unexpected considering the large number of predictors (36) available in this study.

Finally, because of the specific set of predictors that arose from a regression analysis of this job role, it is highly probable that a different profile of EI predictors would arise from testing different job roles. For this reason, and the observation that total EI scores are not effective because they predict so little variance (1%), this study would have to be repeated for each industry or job role that required a specific set of EI predictors for excellent performance.

Summary

This was a transparent and independent study using multiple EI tests and miscellaneous psychological markers to predict success in a specific real-world, industry and job role situation. Standard EI tests were utilised thereby providing a basis for other researchers to replicate this study in the same or other industries. The study found that total EI scores explained an insignificant amount of variance, while the EI subfactors and other factors explained a very significant portion of variance in the performance criterion of up to 52%. However, this brought into question EI theory that proposes that all EI factors are positively related to success. In this regard, this study falsified (Popper, 1963) the theory of EI as a standard intelligence with unidirectional subfactors (e.g., Mayer et al, 1999; Goleman, 1995, 2001b).

As a footnote, the results of this study have been incorporated into a web-based survey that has collected input from 40 prospective job applicants since its implementation in December, 2006. It would be useful to perform a longitudinal predictive validity confirmation study after these new applicants have been actively working in the field as Sales Representatives after a year or so.

Chapter 5

Study 3 (Chapter 4) demonstrated that EI does not conform to the structure of a unified intelligence on the basis that total (global) EI scores were insignificant predictors of the performance criterion; while the EI subfactors, entered as separate predictors, explained a highly significant 38% of the variance in the criterion. Study 3 also found that some EI subfactors were negatively related to high-performance. This falsifies the common assertion that EI subfactors can be simply summed to produce a global EI score (e.g., Bar-On, 1997; Mayer, Salovey & Caruso, 2000c). The summing of predictors with negative and positive directionality appears to dilute the variance explained by the global EI score. These patterns were also found in Study 2 (Chapter 3) confirming that, in this set of studies, EI tests are more akin to multifactorial inventories like the 16-PF (Cattell, 1994), the BFF model of personality (Costa & McCrae, 1985, 1992), and the MMPI (Hathaway, McKinley, Butcher, Dahlstrom, Graham & Tellegen, 1989) than they are to general intelligence with a common *g* factor (e.g., Carroll, 1993).

In an attempt to further generalise these findings, Study 4 was undertaken with the addition of the MSCEIT, a performance-based EI measure (Mayer, Salovey & Caruso, 2002c), and an alternative real-world criterion of customer service performance in a call-centre. Spencer (2001) claims that customer service representatives ($N = 320$) selected on the basis of eight EI competencies demonstrated a 24% productivity increase and a 99% reduction in turnover when compared with average employees. However, EI testing for employee selection is only a recent development, as IQ has traditionally been considered the best predictor of performance for candidates without prior experience in the job role (e.g., Schmidt & Hunter, 1998).

Nonetheless, the narrow range of IQ scores observed in professional employees, under the best conditions, have moderate correlations (< 0.49) with job-performance measures; typically explaining 10% to 30% of the variance in job performance (Matthews, Zeidner & Roberts, 2004). This figure varies somewhat between theorists who propose that under the best conditions, IQ accounts for 25% of the variance (Hunter & Schmidt, 1984, cited in Goleman, 1995), and others who believe that a more realistic figure may be 4% to 10% of the variance (Sternberg, 1996, cited in Goleman, 1995). With IQ only explaining a range of 4% to 30% of job performance variance there is clearly a large percentage of the variance that is unexplained by IQ.

One explanation for other factors such as EI being a superior predictor is that IQ is a given for most high-level job roles. Most professional positions demand a university undergraduate degree, or at least excellent verbal skills, as proof of a certain level of attainment before employers will consider the applicant further. As most university graduates have IQs in the range of 110 to 120 points there is very little variance in this restricted range that can be tapped to establish significant

individual differences (Goleman, 1998). On the other hand, the conative and non-intellective factors (Wechsler, 1950) that Goleman subsumes under EI, have a much broader range than IQ, and thus provide more variance to establish individual differences.

Mackintosh (1998) notes that the use of IQ testing is much less successful in predicting outcomes in university students than with the general public. A large part of the overall theoretical evidence for a common general intelligence (*g*) or IQ factor arises because in a general public sample, some people do poorly on all tests, thus providing substantial variance to establish a robust IQ factor (*g*). He proposes that there is, "no general intelligence, only a lack of general intelligence" or "general stupidity" (Mackintosh, 1998, p. 213). This supports Goleman's assertion that IQ tests only explain limited variance in university-educated samples, because the individuals all perform reasonably well on IQ tests. The university-educated samples produce a limited range of IQ scores with limited variance resulting in low correlations with job performance measures. In this context, the broader ranges of non-intellective factors such as EI, provides a wider range of individual difference variables that can subsequently be more effectively related to job performance metrics.

Goleman (1995) tends to subsume a range of "other characteristics" (p. 34) that explain the remaining variance, under the banner of EI, without acknowledging pre-existing personality and interpersonal differences such as impulsivity and optimism that were established well before EI became a dominant paradigm (e.g., Costa & McCrae, 1985; Mischel, Ebbesen & Zeiss, 1972; Seligman, 1990). Some researchers are quite adamant that EI should not subsume existing constructs (e.g., Matthews et al., 2002, 2004) and should not be presented as "old wine in new bottles" (Salovey, 2006, p. 269). Salovey recommends that all EI studies should include personality, IQ, and other miscellaneous constructs such as Alexithymia (Bagby, Parker & Taylor, 1994, 1994a) to ensure that only new, emotionally-based abilities are labelled EI. Be that as it may, the study reported here used the two foremost EI tests, the MSCEIT and the EQ-i, with their own inherent operationalisation of the EI construct, to explore Goleman's assertion that EI explains a large part of the remaining 80% of job performance variance.

The ability to predict job performance provides opportunities for improving the profit, productivity, and quality of the organization through the selection of superior candidates, and on the other hand, reducing the training and incidental costs arising from the attrition of substandard performance employees. In this regard, the cost of attrition of one employee is equated to one full year's pay (Spencer, 2001).

In a study of 121 companies it was found that two-thirds of the abilities required for effective performance were emotional competencies (Goleman, 1998). In another study of medium complexity job roles such as sales clerks, a performer in the top 1% was found to be 12 times more productive than an employee in the bottom 1% (Cherniss, 2004). Since two thirds of the abilities

predicting superior performance are emotional competencies, EI potentially has a major role to play in predicting employee productivity. Even in areas such as computer programming, where emotional competencies seem to have no place, it was found that teamwork and collaboration abilities enable the top 10% of computer programmers produce three times the output of average performers (Goleman, 1998). For engineers involved in scientific research, EI qualities were four times more important than intelligence in predicting success and prestige (Goleman, 1998)

In an insurance environment, sales agents who scored high in emotional competencies such as self-confidence, empathy, and initiative typically sold policies twice the amount of the average policy (Goleman, 1998). In a cosmetics company, sales trainees that were chosen on the basis of their emotional competence were twice as likely to finish their training as employees not selected according to EI skills.

Many of the claims made by Goleman (1998) and Cherniss (2004) are included in both publications and appear repetitive, and in addition, loosely subsume anything that it is not IQ under the banner of EI. Nevertheless, even if the claims for the advantages accruing to emotional competencies can only be partially achieved, they are compelling and worthy of further examination.

Mayer and his colleagues have claimed that EI had attained the status of a standard intelligence on the basis that the MSCEIT assessed emotional abilities using questions for which there was only one correct answer, consisted of subtasks that were all positively correlated and could be explained by a general factor, total EIQ, and was found to increase with age and experience (Mayer, Salovey, Caruso & Sitarenios, 2001). In an earlier paper Mayer had claimed that his Multifactor Emotional Intelligence Scale (MEIS) had operationalised a global EI factor and had "met the classic criteria of a standard intelligence" (Mayer, Salovey & Caruso, 1997; Mayer, Caruso & Salovey, 1999, p. 267). To claim that a new construct such as EI is an "intelligence" per se, it must also correlate moderately with established measures of general intelligence. In this regard Mayer reported a moderate correlation of (r = .36, p < .001) between the MEIS and verbal IQ (Mayer et al., 1999). Overall, the Mayer et al. definition of EI postulates a global EI factor, comprised of 8 to 12 subfactors that are positively correlated with each other, with sufficient shared variance to theoretically support a single, global EI factor.

Bar-On on the other hand, began developing his test in 1980, well before the advent of the EI concept. His research question was: "Why do some people have better psychological well-being than others?" (Bar-On, 1997, p. 1). Bar-on does not appear as focused as Mayer et al. in establishing EI as a standard intelligence, as his Epilogue in the EQ-i Technical Manual clearly demonstrates (Bar-On, 1997, p. 147):

> Although I do not know precisely how Emotional Intelligence will be defined in the end, I do know that it is part of general intelligence and can be measured along a continuum much like cognitive intelligence.

Although both Bar-On and Mayer provide a single, total EQ score, the predictive validity of this global score requires empirical evaluation as performed in the study reported here. The amount of variance explained by the global EI score was compared with the amount explained by the subfactors for all three EI tests in this study; the MSCEIT, the EQ-i, and the TEIQue. In contrast, earlier multi-dimensional inventories such as the 16-PF or the MMPI recognised the diversity of their subscales and made no attempt to reduce them to one common factor as EI theorists do. It was expected that the data collected from this study would provide some insight into the question of whether EI can be adequately represented by a single global factor.

The sample chosen for this study were telephone claims consultants for a major insurance company. The consultants are involved in negotiation and settling insurance claims under conditions where the customer is often in an emotional state at the time of the call. The ability to communicate clearly and manage the customer's emotional state in these difficult circumstances would appear to be an ideal test of EI abilities of the consultant to manage both self and the customer. The characteristics required for the position include customer focus, communications ability, relationship building, influence and resilience. These skills closely map on to the EI domain; specifically interpersonal and intrapersonal skills (e.g., Bar-On, 1997; Goleman, 1995, 1998).

At this point it would be appropriate to review the introduction and summarise the four main points of interest in this study. First, it is now widely accepted that Wechsler's assertion that intelligence and success includes aspects of personality and drive that go beyond IQ scores. This broadening of the concept represents a more inclusive picture of human performance. Second, IQ scores have limited ability to predict job-performance due to the narrow range observed in university graduates and professionals. Third, there are potential productivity and cost savings benefits accruing from being able to select high-performance candidates using a broad range of tests. In this regard it was hoped that EI would add incremental predictive information over traditional tests to enhance hiring processes. Finally, the assertion by the EI theorists that EI is a unitary intelligence that can be adequately represented by a single global EI score was examined empirically. Overall, it was hoped that the body of knowledge surrounding EI would be significantly enhanced by addressing these four major points.

Summary of aims and hypotheses

The overall aim of this study was to establish the adequacy or otherwise of EI as a predictor of job performance in an organizational setting. This builds on Study 3 (Chapter 4) by introducing

the MSCEIT ability EI test, and investigating an alternate workplace sample to establish whether the previous findings can be generalised. This would confirm or refute the claim that EI can explain a large proportion of the 80% of variance remaining after controlling for IQ. The overall aims can be summarised in four formal hypotheses:

H1: Total EI scores will be significant predictors for the Performance KPI (PKPI) criterion, and better than the IQ predictors.

H2: EI subscales will explain significantly more variance than total EI scores based on the previous studies that indicated EI was not a unified intelligence like general intelligence (*g*).

H3: Total EI scores will discriminate between average (3) and excellent (5) performers as defined by the Performance KPI criterion.

H4: Selected tests will explain a significant amount of variance that will be useful for applicant screening.

Method

One hundred and forty nine Telephone Sales Consultants (124 female – 83%), from a major household and automobile insurance company in Australia took part in this study sponsored by the Human Resources department. The Telephone Sales Consultants process sales and insurance claims enquiries in their day-to-day activities. They completed this battery as an after-hours task and were paid overtime for the 90-minutes it took to answer the questionnaires. Their ages ranged from 18 to 61 years with a mean of 33.5 years (*SD* = 9.8). Ninety four percent of the participants stated that English was their first language. The country of birth of 85% of the participants was Australian or a Western country, leaving 15% from a wide variety of non-English speaking countries. This sample included more Western-born participants than the typical Sydney metropolitan demographic which is comprised of 72% from Western and 28% from non-English speaking backgrounds (ABS, 2001).

Tasks in this Battery of Tests

Three Emotional Intelligence (EI) tests, the MSCEIT, the EQ-i, and the TEIQue-SF, plus a variety of personality, general intelligence, and miscellaneous tests that were expected to correlate with employee performance were administered. In addition, a report containing key performance indicators (KPIs) in eight areas was provides by individual managers for each employee. A summary of the families of tasks in the experimental battery is shown below in Table 1.

Table 24

Battery of tasks performed by each participant in this study with the criterion of the performance of the individual as determined by the manager as the final metric.

Task	Type	Author/Source
Emotional Intelligence tests		
1. MSCEIT	ability EI	Mayer et al. (2002)
2. EQ-i	mixed EI	Bar-On (1997)
3. TEIQue	self-report EI	Petrides et al. (2003)
BFF Personality Inventory		
4. Mini-Markers	self-report	Saucier (1994)
General Intelligence		
5. Word reasoning	verbal (gc)	Stankov (2000)
6. Numerical reasoning	numeric (gf)	Stankov (2000)
Miscellaneous tests		
7. Motivation	self-report	Various
8. Optimism	self-report	Various
9. Impulse control	self-report	Tellegen MPQ (1982)
10. Empathy	self-report	Davis (1980)
Criterion		
11. Performance KPI measures	manager report	HR department

Biographical Details

The first questionnaire containing the TEIQue collected the country of birth and first-language of the participant, while the MSCEIT and the EQ-i provided the age and gender of the participants.

Emotional Intelligence tests

Mayer, Salovey and Caruso Emotional Intelligence Test (MSCEIT). The MSCEIT Version 2 (Mayer et al., 2002c) is a 141-items test divided into eight sections that produce eight emotional ability Task scores: Faces, Pictures, Sensations, Facilitation, Changes, Blends, Emotional Management and Emotional Relations. The eight subfactors are paired to produce four Branch scores: Perceiving emotions, Facilitating thought, Understanding emotions and Managing emotions. These are further combined to produce two Area Scores: Experiential EI and Strategic EI quotients (EIQ). Finally, the two Area scores are combined to produce the Total EIQ, the overall score.

The MSCEIT is consensus scored, that is, the database of responses collected to date have been analysed, and the percentage of responses for each item have been used to provide the scoring

keys for participants. The scoring keys are proprietary and the test can only be scored through the supplier, Multi-Health Systems (MHS).

Bar-On Emotional Intelligence Inventory (EQ-i). The EQ-i is a 133-item self-report measure consisting of a mixed array of 15 trait subfactors such as Assertiveness and Optimism that are combined to produce a Total Emotional Quotient score, the Total EQ score. The items consist statements such as: "I like everyone I meet" that are rated on a Likert scale where, 1 = "*Very seldom true of me*" and 5 = "*Very often true of me*." The fifteen factorial components are combined to produce five composite scales: Intrapersonal EQ, Interpersonal EQ, Stress management EQ, Adaptability EQ, and General Mood EQ.

Although the EQ-i scoring keys are proprietary and all tests must be scored through MHS, the scoring is more straightforward than the MSCEIT and can be interpreted further if the researcher wishes to explore the results in detail. The actual participant responses for each item are supplied with the total scores for the EQ-i and the MSCEIT to enable further analysis if required.

Trait Emotional Intelligence Questionnaire – Short Form (TEIQue-SF). The TEIQue-SF is a 30-item, short-form version of the 144 item TEIQue, Version 1 (Petrides, Perez & Furnham, 2003, 2004) . The TEIQue-SF was designed to yield a global EI score and does not provide the 15-subfactors that are available in the 144-item TEIQue. The TEIQue-SF is built on a 4-factor theoretical structure: Emotional Well-being, Self-control, Emotionality and Sociability.

The questions such as, "I usually find it difficult to regulate my emotions" are responded to on a 7-point Likert scale ranging from: 1 = "*Completely disagree*" to 7 = "*Completely agree*." The scoring key for the global trait EI score is provided by the authors on the questionnaire documentation, and the scoring key for the four factors was obtained from the author K. V. Petrides (personal communication, March 1, 2006). For simplicity, this paper will omit the "-SF" suffix and simply refer to the test as the TEIQue throughout this document.

Big-Five Factor Personality Inventory (BFF)

The BFF personality inventory employed in this study was Mini-Markers; a brief version of Goldberg's unipolar Big-five markers (Saucier, 1994). This 40-item inventory demonstrates high alpha reliabilities on all five factors: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism in an easy to administer test. For simplicity, the abbreviations of O, C, E, A and N will be used throughout this chapter to describe the five factors above.

The instructions for questionnaire asks the participant, "*How accurately can you describe yourself*?" It then presents 40 personality adjectives such as, "*Bold*," "*Shy*," and "*Warm*," that the participant rates on a 9-point Likert scale ranging from, 1 = "*Extremely inaccurate*" to 9 = "*Extremely accurate*." The scoring key and factorial structure is provided with the documentation to enable the experimenter to score results. In this study, the Likert scale was reduced to 5-points, still

using the same endpoints as described above, to enable comparisons with a larger sample collected at the University using computer readable answer sheets that only had five response options. The reduction of the Likert scale to 5-points did not appear to have an adverse effect on alpha reliabilities.

General intelligence markers (IQ)

While the markers for crystallised intelligence (Gc) and fluid intelligence (Gf) did not constitute a comprehensive IQ test, it is useful to compare them with EI test results. The general intelligence dimension of individual differences will be referred to as IQ for brevity throughout this study, although it is acknowledged that this is not a complete IQ test.

Word Reasoning (Word). The Word reasoning task was derived from 20-items from the Esoteric Analogies test in the Gf/Gc Quickie Test Battery (Stankov, 1997). The participant is presented with a sentence with three key words such as: "*Horse is to Animal as Chair is to*:" and four alternate answers to choose from, such as: "*Furniture, Leg, Person, Table.*" The participant's total score out of 20 is simply the number of items that they answer correctly. The Word task is a marker for verbal skills or crystallised intelligence (*gc*) (Carroll, 1993).

Numeric Reasoning(Odd-one-out). The Odd-one-out task consists of a 10-item subset of questions taken from the Gf/Gc Quickie Test Battery (Stankov, 1997; Stankov & Roberts, 2001). The participant is asked to determine which sequence of four numbers is the odd-one-out, for example:

"*7-4-8-9,*" "*6-5-9-2,*" "*5-4-8-3,*" "*9-4-8-5,*" "*6-4-8-2.*"

The rules change for each of the ten items in this task, thus it is a measure of ability to use reasoning to see abstract patterns, which is the defining property of fluid intelligence (*gf*) (Carroll, 1993).

Miscellaneous inventories

Motivation. The Motivation test items, designed to test motivation to achieve and succeed, were assembled from a variety of public domain sources. Items were reverse scored and counterbalanced to avoid any effects from a tendency to respond in the affirmative. A five-point Likert scale ranging from: 1 = "*Never*" to 5 = "*Always*" was utilised. Examples of motivation items are as follows:

"*I avoid positions of leadership (r)* "

"*My level of motivation to achieve is high*"

"*I seek a quiet life and avoid challenges (r)*"

The responses were reversed where applicable *(r)* and summed to provide a total Motivation score.

Optimism. The Optimism scale, designed to establish the participant's score on the optimism/pessimism dimension, was assembled from a variety of public-domain sources. The

20-items were reverse scored and counterbalanced to avoid response bias, and a 7-point Likert scale was utilised with: 1 = "*Strongly disagree*" and 7 = "*Strongly agree*" as the extreme anchor points. Examples of Optimism items are as follows:

"*I am optimistic about the future*"

"*The future looks bleak (r)*"

"*Good things happen to me*"

The responses were reversed where applicable and summed to provide a total Optimism score.

Impulse control.

The 16-items that comprise the Impulse control inventory were taken from Tellegen's Multidimensional personality questionnaire (MPQ) (Tellegen, 1982). Items were scored on a 5-point Likert scale where: 1 = "*Never true*" to 5 = "*Always true*." Examples of Impulse control items are as follows:

"*I often act without thinking (r)*"

"*I usually make up my mind through careful reasoning*"

"*I often act on the spur of the moment (r)*"

The responses were reversed where applicable and summed to provide a total Impulse control score.

Empathy. The 13 Empathy items were derived from the 27-item Interpersonal Reactivity Index (Davis, 1980) with items such as, "*I really get involved with the feelings of the characters in a novel*" omitted because they do not apply to corporate activities. Items were scored on 5-point Likert scales where: 1 = "*Strongly disagree*" to 5 = "*Strongly agree*." Examples of Empathy items are as follows:

"*I don't feel very sorry for other people when they are having problems (r)*"

"*I believe that there are two sides to every question and try to look at them both*"

"*I sometimes find it difficult to see things from the other person's point of view (r)*"

The responses were reversed where applicable and summed to provide a total Empathy score.

Manager report – The criterion

The Manager's report was designed by the HR department to be used as the real-world criterion for individual performance. It took the form of eight, Key Performance Indicators (KPIs) for each participant. The KPIs were rated on a 5-point Likert scale where: 1 = "*Poor*," 2 = "*Below average*," 3 = "*Average*," 4 = "*Above average*," and 5 = "*Excellent*." The eight KPIs are:

1. Decision Making
2. Customer Focus
3. Telephone Communication
4. Develops and Maintains Relationships
5. Persuasiveness/Influencing

6. Quality Focus
7. Resilience
8. Performance KPIs (PKPI)

The KPI ratings one through seven consisted of the manager's subjective performance rating for each individual, while number eight, the performance KPI (PKPI) was an objective measurement of yearly performance metrics, such as average phone time, that were not a subjective evaluation by the manager. The PKPI scores were correlated with the survey results to determine the predictive properties of each of the tests: BFF personality, IQ, miscellaneous inventories, and the 27 subscales of the EI tests, making a total of 38 subfactors.

Apparatus

All tasks for this study were internet-based using IBM-compatible computers located on the client's premises, running Windows™ XP. All surveys were designed to be hosted on a university server and invoked on a PC using the Microsoft Internet Explorer.

Test Procedure

Participants took the tests after-hours, away from their own desks in the work environment, and thus were not interrupted by telephone calls and day-to-day business activities. The battery of tests consisting of the three on-line questionnaires typically took 90-minutes to complete.

Results

Descriptive Statistics

The following table lists the tests used in this battery, together with normative data and alpha reliabilities where available.

Table 25

Descriptive statistics for the battery of tests (N=103), with coefficient alpha, and standardised values from the original test-developers where available.

	This Study			Standardised Data		
Test	Mean	SD	Alpha	Mean	SD	Alpha
MSCEIT Total EIQ	95.10	13.20	(1)	100	15	.93
EQ-i Total EQ	93.76	13.39	(1)	100	15	.80
TEIQUE Total EI	5.07	.76	.89	5.07	.65	.86
Openness	28.79	5.39	.74	30.9	4.6	.78
Conscientiousness	32.09	4.69	.74	28.4	5.1	.78
Extraversion	27.46	5.56	.73	26.9	5.4	.81
Agreeableness	33.51	4.45	.71	32.6	4.4	.81
Neuroticism	18.14	5.09	.75	20.4	4.8	.79
Word reasoning score (Gc)	.69	.16	.73	0.80	.11	.71
Numeric reasoning score (Gf)	.75	.19	.61	0.73	.27	.87
Motivation	3.27	.51	.66	3.36	.48	.69
Optimism	4.83	.83	.88	(2)	-	-
Impulse control	3.51	.55	.83	3.33	.57	.84
Empathy	3.75	.64	.78	3.76	.61	.77
Performance KPI (PKPI)	3.81	.86	(3)	-	-	-

Notes: 1) Alpha reliabilities were not provided by MHS with the scored sample. 2) This is a new test and there is no standard data. 3) The PKPI is a single rating by the manager and alpha reliabilities cannot be calculated.

All skew and kurtosis statistics were within the range of plus one to minus one; the maximum statistic being 0.68, indicating that the data did not require transformation before further analysis (Dekker, 2005).

Skew in the criterion (PKPI)

The performance KPI criterion was negatively skewed, with most of the participants falling in the 3 to 5 range. This can be seen in Figure 8.

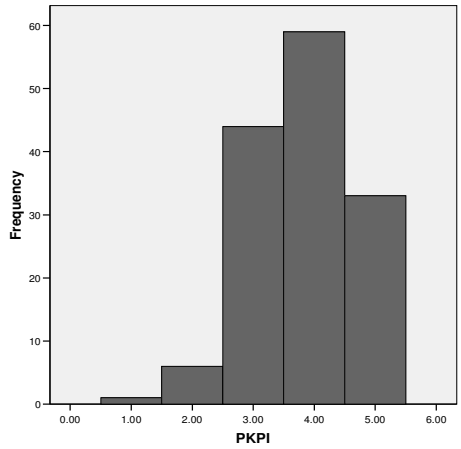


Figure 8

The histogram above shows the number of participants in each category of performance KPIs (PKPI) causing skew because there are only seven participants with scores of 1 and 2.

Although the skew statistic for PKPI was only -.31, and thus did not exceed plus and minus 1 as is the recommended limit for skew (Dekker, 2005), it nevertheless produced an effect where seven, poor and below average, participants had an excessive influence on the statistics. The seven participants with 1 and 2 PKPI ratings were included in the correlation matrices that follow, but excluded from the regression analyses.

With the 1 and 2 scorers removed, the 3 to 5 distribution approximates normal, however it only has three categories, 3, 4 and 5. For this reason Ordinal Logistical Regression was utilised for regression analysis.

Correlation Matrix

Table 26

Relationships between major variables in the test battery.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. MSCEIT Total EIQ	1													
2. EQ-i Total EQ	.22*	1												
3. TEIQUE Total EI	.23*	.81**	1											
4. Openness	.25**	.07	.14	1										
5. Conscientiousness	.16	.39**	.35**	.09	1									
6. Extraversion	.11	.37**	.36**	.27**	.14	1								
7. Agreeableness	.26**	.22*	.23**	.19*	.26**	-.08	1							
8. Neuroticism	-.13	-.47**	-.49**	.02	-.27**	-.10	-.35**	1						
9. Word reasoning score	.39**	.09	.11	.10	-.08	-.04	-.03	-.11	1					
10. Numeric reasoning score	.08	.07	.21*	.00	-.02	.03	.04	-.01	.25**	1				
11. Motivation	.08	.28**	.30**	.32**	.15	.40**	-.04	.01	-.05	.03	1			
12. Optimism	.10	.62**	.71**	.08	.30**	.24**	.13	-.40**	.01	.09	.33**	1		
13. Impulse control	.05	.20*	.21**	-.02	.48**	-.12	.11	-.18*	.10	.01	.03	.29**	1	
14. Empathy	.31**	.46**	.42**	.20*	.15	-.02	.59**	-.27**	.14	.03	.15	.30**	.16*	1
15. Performance KPIs	-.13	.10	.10	-.03	.20*	-.08	.08	-.09	.13	.24**	.00	.12	.20*	.12

Note: ** indicates that correlation is significant at the 0.01 level (2-tailed), * indicates significance at the 0.05 level (2-tailed).

The correlation matrix (Table 26) indicates that only three tests correlate significantly ($p < .05$) with the performance KPIs (PKPI) criterion. The first was the Conscientiousness factor from BFF personality theory, followed by the fluid intelligence marker (*gf*) from the odd-one-out test, and finally the impulse control scale from the miscellaneous tests. Neither the MSCEIT, the EQ-i, or the TEIQue global scores correlate significantly with the PKPI criterion, and all exhibit very low correlations of -.13, .10, and .10 respectively. It is noteworthy that the MSCEIT total EIQ is negatively correlated with performance, with the interpretation that higher PKPI performers score lower on the MSCEIT. The issue of lack of predictive validity of the EI tests will be explored in more detail in the regression analysis.

While the total EI scores for all three tests do not reach significance in predicting performance on the PKPI criterion, MSCEIT Facilitation and EQ-i Social responsibility do reach significance (Table 27). Surprisingly, the MSCEIT Facilitation task score is negatively correlated with performance on the PKPI criterion (Table 27). Overall, five out of the eight (63%) MSCEIT task scores are negatively correlated with the PKPI criterion.

The EQ-i only has one factor, Social Responsibility, that is significantly and positively correlated with the PKPI criterion. The EQ-i has two out of the 15 factors (13%), Self Regard and Assertiveness that are negatively, albeit insignificantly, correlated with the PKPI criterion.

Relationship between the Performance KPI criterion and EI subscales

Table 27

Correlations between the MSCEIT, EQ-i, and TEIQue variables and the job-performance KPI criterion.

EI Variable	Performance KPIs
MSCEIT Total EIQ	-.13
MSCEIT Faces	-.15
MSCEIT Facilitation	-.21*
MSCEIT Changes	-.07
MSCEIT Emotion Management	.15
MSCEIT Pictures	-.12
MSCEIT Sensations	-.13
MSCEIT Blends	.10
MSCEIT Emotional Relations	.12
EQ-i Total EQ Score	.10
EQ-i Self regard	-.09
EQ-i Emotional self awareness	.08
EQ-i Assertiveness	-.04
EQ-i Independence	.10
EQ-i Self actualization	.14
EQ-i Empathy	.16
EQ-i Social responsibility	.22*
EQ-i Interpersonal relationship	.08
EQ-i Stress tolerance	.10
EQ-i Impulse control	.08
EQ-i Reality testing	.08
EQ-i Flexibility	.10
EQ-i Problem solving	.03
EQ-i Optimism	.11
EQ-i Happiness	.01
TEIQue Total EI	.10
TEIQue Well-being	.11
TEIQue Self-control	.04
TEIQue Emotionality	.12
TEIQue Sociability	.05

Note: * indicates that correlation is significant at the 0.05 level (2-tailed).

H1: Total EI scores as better predictors of the PKPI criterion than IQ

Hypothesis one predicted that total EI scores from the MSCEIT, the EQ-i and the TEIQue would be significant predictors of the Performance KPI (PKPI) criterion. To explore this

hypothesis, a series of analyses were conducted, first with the total EI scores from each test, then the subfactors from each test. In addition, a regression analysis (Model 9, Table 28) was conducted for the general intelligence (IQ) markers: verbal reasoning (Word) and numeric reasoning (Odd-one-out). This is used to make comparisons with the variance explained by IQ to that explained by the EI tests (e.g., Goleman, 1995). This shows that 8% of the variance in the PKPI criterion can be explained by IQ, although Numeric reasoning was the only significant predictor.

The correlations between the PKPI criterion and the MSCEIT, the EQ-i, and the TEIQue were -0.13, 0.10 and 0.10 respectively (see Table 26). The MSCEIT scores indicated a negative relationship where higher performers (in terms of higher PKPIs) obtained lower scores on the MSCEIT. This result is contrary to the claim that EI predicts outstanding job performance (Goleman, 1998). The EQ-i and the TEIQue both exhibited a positive relationship with the PKPI criterion but the small correlation represents a variance explained by each of only 1%. In comparison, the variance in the PKPI criterion explained by IQ is 8% (Table 28, model 9).

Thus, hypothesis one (H1) is rejected because IQ explained significantly more variance than total EI scores.

H2: EI subscales will explain more variance than total EI scores

To examine the predictive properties of each EI test on the Performance KPI (PKPI) criterion, an ordinal logistic regression analysis was conducted to determine the variance explained by: (i) the Total (global) EI scores, and (ii), the subfactors of the EI tests. The results of this analysis are tabulated in Table 28.

Table 28

Regression analysis of the variance explained by the three EI tests, and their subscales, with the performance KPI (PKPI) as the criterion.

Model	Predictors for Performance KPI (PKPI)	Parameter Estimate	*p* -value	Overall model fit *p* - value	Total variance explained R^2
1.	MSCEIT Total EIQ	-.01	.54	.53	0.4%
	MSCEIT Branch Scores (4)				
	Perceiving emotions	-.02	.13		
2.	Facilitating thought	-.02	.16	.05	1%
	Understanding emotions	.00	.98		
	Managing emotions	.04	.02		
	MSCEIT Task Scores (8)				
	Faces	-.01	.47		
	Pictures	-.02	.16		
	Sensations	-.02	.31		
3.	Facilitation	-.02	.24	.05	15%
	Changes	-.03	.15		
	Blends	.02	.28		
	Emotion management	.01	.52		
	Emotional relations	.04	.06		
4.	EQ-i Total EQ	.02	.09	.09	3%
	EQ-i Subscales (5)				
	Intrapersonal	-.04	.09		
5.	Interpersonal	.04	.01	.05	10%
	Stress management	.02	.37		
	Adaptability	.02	.35		
	General mood	.01	.65		

	EQ-i Factors (15)				
	Self-regard	-.05	.02		
	Self-awareness	-.02	.31		
	Assertiveness	-.06	.01		
	Independence	.02	.23		
	Self-actualization	.05	.03		
	Empathy	.00	.91		
	Social responsibility	.02	.27		
6.	Interpersonal relationship	.03	.23	.01	27%
	Stress tolerance	.04	.10		
	Impulse control	.00	.93		
	Reality testing	.01	.53		
	Flexibility	.01	.67		
	Problem solving	-.02	.33		
	Optimism	.02	.58		
	Happiness	.01	.57		
7.	TEIQue global EI score	.13	.55	.54	0.3%
	TEIQue factors				
	Well-being	.00	1.00		
8.	Self-control	-.02	.94	.97	0.5%
	Emotionality	.12	.60		
	Sociability	.04	.86		
	General Intelligence (IQ)				
9.	Word reasoning (*gc*)	1.24	.23	.01	8%
	Numeric reasoning (*gf*)	2.33	.03		

Stepwise regression analysis

Stepwise regression analysis using the *backward* elimination method was conducted to determine the best predictors from the MSCEIT, EQ-i, and the TEIQue EI tests, the personality (BFF) and the miscellaneous tests to explain the maximum amount of variance in the PKPI criterion. Table 29 lists the EI subfactors and individual tests that were found to be significant predictors of the PKPI criterion.

Table 29

Ordinal Logistic Regression analysis yielded the following values for significant predictors of the Performance KPI (PKPI) criterion. Note that all predictors are significant at $p < .05$.

Model	Predictors for the Performance KPI (PKPI) criterion	Parameter Estimates	Parameter *p* - value	Overall Model Fit *p* – value	Total variance explained R^2
	MSCEIT Facilitation	-.03	.040		
	MSCEIT Sensations	-.05	.018		
	MSCEIT Emotional relations	.05	.012		
	EQ-i Self regard	-.09	.001		
	EQ-i Assertiveness	-.09	.001		
	EQ-i Self-actualization	.09	.002		
1.	EQ-i Interpersonal-relationship	.10	.001	<.0001	50%
	EQ-i Stress tolerance	.09	.001		
	EQ-i Problem solving	-.07	.007		
	IQ Word reasoning score (Gc)	5.51	.001		
	IQ Numeric Reasoning (Gf)	3.38	.013		
	BFF Conscientiousness	.15	.018		
	TEIQue Emotionality	-.73	.041		

The regression analysis in Table 29 indicates that 50% of the variance in the PKPI criterion can be explained by a mixture of EI factors, BFF personality, and miscellaneous tests. All predictors are significant at $p < .05$ and the overall model is an excellent fit at $p < .0001$. The R^2 variance shown in Table 29 is the Nagelkerke "Pseudo R-Squared" statistic provided in the SPSS output of an ordinal logistic regression analysis. SPSS also provides the "Cox and Snell" pseudo R^2 however this statistic errs on the low side and never reaches a theoretical maximum of 1; thus the Nagelkerke statistic is preferred (Field, 2005).

In summary, MSCEIT EI subfactors explained 15% (Table 28, model 3), and the EQ-i explained 27% (Table 28, model 6) of the variance in the PKPI criterion. This is in contrast to the total EI scores explaining 3% or less of the variance (models 1 and 4). The difference between 3% of the variance explained by the EQ-i (Table 28, model 4) and 1% from the correlation matrix ($r = .10$, Table 26) is probably due to dropping 1 and 2 from the PKPI criterion and only using values 3, 4 and 5, and variations introduced by the ordinal logistic regression analysis. Nonetheless, the difference in variance explained by the total EI scores, 3% versus the 15% to 27% of the variance explained by the subfactors clearly indicates that the subfactors explain significantly more variance than the global EI scores.

Thus, hypothesis two (H2), that EI subscales will explain more variance in the PKPI criterion than total EI scores, is accepted on the basis of these results.

Tests that predict the difference between average and excellent performers

The following analysis shown in Table 30 applies to hypotheses three and four. To determine which tests discriminate between average and excellent performers an independent samples *t*-test was conducted on all 38 tests and EI subfactors utilised in the battery. Only those tests with a *p*-value of less than 0.05 exhibit significant differences between average and excellent performers.

Table 30

Tests that discriminate significantly between Average (3) and Excellent (5) performers as defined by the Performance KPI (PKPI) criterion.

Variable	*t*	*df*	*p* - value (2-tailed)	Mean Difference	Mean score PKPI = 3	Mean score PKPI = 5
MSCEIT Total EIQ	-0.66	61	.51	-2.09	95.97	93.88
EQ-i Total EQ	1.79	64	.08	5.87	91.00	96.88
TEIQue global EI score	0.65	68	.52	0.12	5.06	5.18
MSCEIT Emotional Relations	2.08	63	.04	6.10	91.30	97.40
EQ-i Self actualization	2.43	64	.02	7.42	88.26	95.69
EQ-i Empathy	2.53	64	.01	9.53	93.00	102.53
EQ-i Social responsibility	3.30	64	.002	10.16	93.09	103.25
IQ Numeric reasoning (*gf*)	3.00	68	.001	0.12	0.69	0.81
BFF Conscientiousness	2.99	68	.004	3.07	30.51	33.58

Note that all subfactor differences are statistically significant at $p < .05$, and are all in the theoretically predicted direction; with *Excellent* performers (5) scoring higher on all six measures, except the MSCEIT Total EIQ. Clearly, numeric reasoning ($p = .001$) is the most discriminative test, followed by EQ-i Social responsibility ($p = .002$), and BFF Conscientiousness ($p = .004$). This data is shown in graphical format in Figure 9.

H3: Total EI scores will discriminate between average and excellent performers

The results shown in Table 30 indicate that none of the global EI test scores for the MSCEIT, EQ-i or the TEIQue reached significance in a *t*-test between average and excellent performers. The EQ-i came the closest of any EI test in being able to predict average from excellent performers ($p = .08$) but failed to reach significance ($p < .05$).

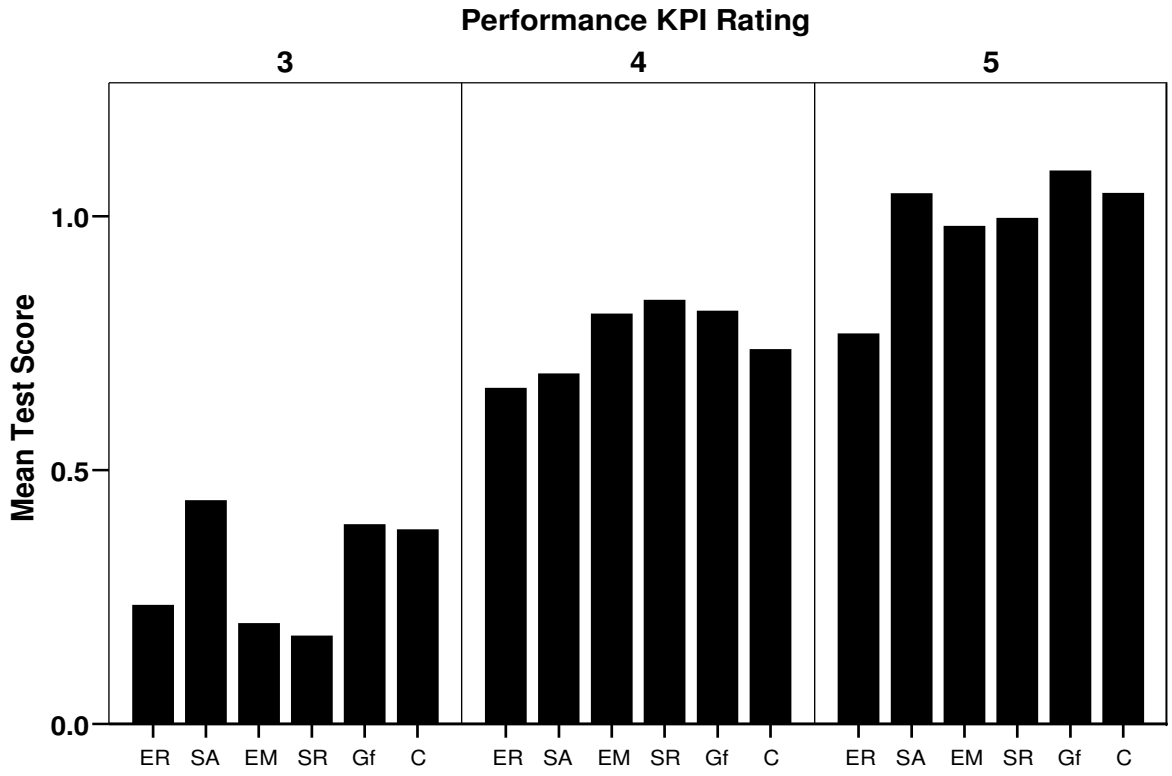


Figure 9

Graphical representation of Table 30 showing how mean test scores of selected subfactors increase as performance KPI ratings of participants go from 3 to 5. The abbreviations are as follows: ER = Emotional relations, SA = Self actualization, EM = Empathy, SR = Social responsibility, Gf = Numeric reasoning and C = Conscientiousness.

Thus, hypothesis three (H3) is rejected; total EI scores do not discriminate between average (3) and excellent (5) performers as measured by the PKPI criterion.

H4: Selected tests will explain significant variance in the PKPI criterion

Regression analysis output shown in model 7 in Table 29 demonstrated that 13 selected tests, EI subfactors and miscellaneous tests, can explain up to 50% of the variance in the PKPI criterion. By creating a short selection battery that included these tests, an overall score could be produced that would provide a useful metric for selecting potential candidates for this industry.

An alternative approach for selecting tests for applicant screening is by using a *t*-test to determine which subtests show statistically significant differences between the mean scores of average versus excellent performers (Table 30). The six subfactors are also shown graphically in Figure 9, demonstrating a gradual increase in the average of all six subfactors as performance ratings (PKPI) move from average (3) to excellent (5).

Hypothesis four (H4) is therefore accepted; selected tests explain a significant amount of variance in the PKPI criterion ($p < .0001$).

Discussion

The four hypotheses for this study were: (H1) that global EI scores will be significant predictors of the job-performance criterion, and more predictive than IQ; (H2) that EI is not a unified construct like general intelligence (*g*), and therefore, EI subfactors will explain significantly more variance than global EI scores; (H3) that global EI scores will discriminate between average and excellent job-performance, and; (H4) that selected tests will provide a suitable model for screening applicants. The following discussion will focus on the implications and findings in the light of these aims and hypotheses.

General observations

Although the managers reported on eight criteria that were initially planned for use in the analysis, only the performance KPI (PKPI) demonstrated significant and consistent correlations with the tests in the battery. This appeared to be due to the PKPI being derived from objective measures of performance while the other seven relied on a subjective evaluation by the managers. Results indicated that the only reliable criterion available for use in this study was the performance KPI measure (PKPI).

It was interesting to note that the MSCEIT performance EI test results had small correlations of 0.22 and 0.23 respectively with the self-report EI tests; the EQ-i and the TEIQue (see Table 26). This replicates the result in a previous study where the MSCEIT and the EQ-i were only correlated at 0.21 (Brackett & Mayer, 2003). The performance and self-report EI tests thus only overlap with 5% of shared variance, which typically would indicate that they are divergent constructs. Although no clear correlations cutoffs are given for convergent validity (e.g., Anastasi & Urbina, 1997) it would be appropriate to expect different tests of the same construct to have moderate to large correlation coefficients (greater than 0.30). For example, the EQ-i and the TEIQue global scores demonstrated excellent convergent validity with a correlation of $r = .81$, $p < .01$ (see Table 26). Although theorists have proposed that there are two types of EI, performance and trait (e.g., Petrides & Furnham, 2001), there is little justification to support the ongoing confounding of these two dissimilar constructs in the general EI literature.

Total EI scores as predictors of performance (H1)

Hypothesis one was that total (global) EI scores would be significant predictors for the Performance KPI (PKPI) criterion, and better than the IQ markers as predictors.

The total EI scores for the MSCEIT, EQ-i and the TEIQue utilised in this study were found to be insignificantly correlated (-.13, .10, .10) with the PKPI criterion and consequently failed to support the thesis that EI is the major differentiator for excellent performance (e.g., Goleman, 1995, 1998, 2002; Cherniss, 2001, 2004; Caruso, Salovey & Mayer, 2003). In fact, the MSCEIT total EIQ score is negatively correlated with the performance criterion (PKPI), albeit insignificantly, demonstrating that the MSCEIT operationalisation of total EI as a predictor of superior performance was not supported empirically in this study. The EQ-i total EQ score also had a very small and insignificant correlation with the PKPI criterion; as did the TEIQue total EI score.

The IQ marker *Word reasoning* had a small and insignificant correlation with the PKPI criterion, but the *Numeric reasoning* IQ marker was a significant predictor ($p < .01$) of the PKPI criterion.

Thus, hypothesis one was rejected on the basis that none of the global EI scores predicted the performance criterion, and a traditional IQ marker, Numeric reasoning was a superior predictor of the PKPI criterion. This is an unexpected result considering the claim was that EI is a more important predictor of success at work than IQ (Goleman, 1995). A university degree is not a requirement for a job as a Telephone Sales Consultant so one possibility was that the range of IQ scores was broader than a university-educated sample would exhibit. However, the results for Numeric reasoning indicate this sample scored slightly higher than the university sample, and the variation in scores (*SD*) was smaller than the university sample. Thus, it is unclear why the Numeric reasoning score was more predictive than all other tests; however, it was clearly a better predictor than global EI scores. Contrary to Goleman's assertions, this result supports the claim that general mental ability is "the most valid predictor of future performance" (Schmidt & Hunter, 1998, p. 262).

EI as a unified intelligence (H2)

MSCEIT. Hypothesis two anticipated that EI subscales would explain significantly more variance than total EI scores because previous studies (Chapters 2, 3 and 4) found that the EI construct was not a unified intelligence; that is, where subfactors shared significant common variance (like psychometric *g*).

Table 28, model 3, shows that the eight MSCEIT task level scores explained a significant 15% of variance ($p = .05$), compared with the MSCEIT total EIQ score explaining an insignificant 0.4% of variance ($p = .54$) in the PKPI criterion. Thus, hypothesis 2 is supported for the MSCEIT. A further concern is that five out of eight of the task-level scores were negatively related with superior performance as measured by the PKPI criterion. The MSCEIT total EIQ score was also negatively correlated with the performance criterion indicating that superior performers obtain lower scores on the MSCEIT total EIQ than do average performers.

This result, indicating that superior employees score lower on the MSCEIT is contradictory to the general premise of EI (e.g., Goleman, 1995; Cherniss, 2000; Caruso et al., 2003). The five predictors in the regression analysis consisting of Faces, Pictures, Sensations, Facilitation and Changes, also have negative parameter signs (Table 28 , model 3) indicating that with all other variables held constant, lower scores on these five predictors are associated with superior PKPI performance. While it may be argued that these negative predictors are not statistically significant, it does raise doubts regarding the proposition that superior performing individuals score high on all EI factors (e.g., Goleman, 1998). Nonetheless, results from the MSCEIT support hypothesis two in line with previous studies (Chapters 2, 3 and 4), because the subscales explain significantly more variance than the MSCEIT total EIQ score; once again refuting the assertion that EI is a unified intelligence with a common EI factor.

EQ-i. Bar-On's EQ-i performed slightly better with total EQ scores explaining an insignificant 3% of the variance of the PKPI criterion. However, the five subscales explained 10% of the variance, and the 15 factors explained a significant 27% of the variance of the PKPI criterion (see Table 28). Thus, hypothesis two is supported for the EQ-i.

Four of the 15 subscales have negative partial correlations with the criterion indicating that higher PKPI scores are associated with lower scores on the Self-regard, Self-awareness, Assertiveness and the Problem-solving factors, when all other variables are held constant. Once again, the inability of a global EI factor to explain significant variance in the performance criterion, whilst the subfactors can, refutes the claim that EI is a unified intelligence with a common EI factor.

TEIQue. The TEIQue global EI score only explained 0.3% of the variance, while the four separate subfactors explained 0.5% of the variance in the PKPI criterion (Table 28). Thus, hypothesis two is nominally supported for the TEIQue; however, the results are really a null outcome because the variance explained in both cases was trivial, and neither regression model was statistically significant.

In summary, none of the global EI scores explained a significant amount of variance, while both the MSCEIT and EQ-i subfactors explained significant variance in the PKPI criterion. Some predictors appeared to have negative beta weights despite EI theory that assumes that all EI factors are positively related to success criteria. The overall results indicate that EI is not a unified intelligence as proposed by theorists (e.g., Goleman, 1995; Mayer, 2001) and to a lesser extent by Bar-On (1997). In this regard the EI tests utilised in this study demonstrated the properties of a divergent, multifactorial collection of constructs such as the 16-PF (Cattell, 1994), the MMPI (Hathaway et al., 1989), or a BFF personality inventory (Costa & McCrae, 1992; McCrae and John, 1992). While large samples may demonstrate a positive manifold indicative of a unified

intelligence, this does not guarantee that all subfactors will have positive correlations with a workplace performance criterion such as the PKPI.

Total EI scores as discriminators of performers (H3)

Hypothesis three proposed that total EI scores would discriminate between average (3) and excellent (5) performers on the Performance KPI criterion. Table 30 lists the tests that demonstrated significantly different mean scores between average (3) and excellent (5) performers. The total EI scores do not reach statistical significance in the *t*-test comparison between average and excellent performers. The graphical representation in Figure 9 demonstrates a smooth progression and a positive increase in selected factors as ratings proceed from average (3) through very-good (4) to excellent (5) rating in the performance criterion.

Total EI scores for the MSCEIT, EQ-i, and the TEIQue did not significantly discriminate between average and excellent performers. Thus, hypothesis three is rejected on the basis that total EI scores do not discriminate between average and excellent performers. This is a robust finding considering that three EI tests returned the same result.

Selected tests as a model (H4)

Stepwise regression (backward) yielded 13 significant predictors ($p < .05$) that explained approximately half the variance of the PKPI criterion (Table 29). These predictors were then utilised in an ordinal logistic regression analysis to determine parameter estimates (*B*-weights) and *p*-values. All predictors were significant at $p < .05$, and the total model was highly significant ($p < .0001$). In four recent studies using the EQ-i it was found that selected factors explained, 24%, 26%, 28%, and 30% of the variance respectively in the criterion (Bar-on, Handley & Fund, 2006). By comparison, the ability to explain 50% of the variance demonstrated in this study surpasses typical levels achieved in similar studies. Model 6 in Table 28 shows that the EQ-i subfactors explained 27% of the variance in this study; quantitatively similar to the Bar-On studies listed above. However, when other factors such as BFF Conscientiousness and other EI subfactors (from the MSCEIT and TEIQue) were added to the predictors, the variance-explained rose to 50% (Table 29). This provides unequivocal evidence that there are factors beyond specific EI implementations (EQ-i) that explain significant incremental variance; 23% in this example. However, the overhead of having to include three EI tests (MSCEIT, EQ-i and the TEIQue) in one battery to explain 50% variance indicates that single EI tests are deficient; they do not include all possible EI factors.

Overall, hypothesis four was accepted on the basis that the ability to explain 50% of the variance provides a useful starting point for applicant screening tests. Further study would of course be required to confirm predictive validity of these variables.

Limitations of this study

The number of participants (149) was a reasonable sample for this study, although missing data often reduced the effective number to 107 due to listwise deletions. The loss of data was mainly due to the problem of MHS not providing a facility where the initial test battery could be seamlessly linked to the EQ-i and the MSCEIT tests. Psychological test companies appear to assume that their test is the only important one to be administered, and do not make provision for linkage to other tests in a battery. Thus, most of the data loss appears to be due to participants being confused about manually entering URLs, and consequently missing one of the tests; the MSCEIT or EQ-i.

The manager ratings on a 1 to 5 Likert scale proved to have too few steps to provide maximum differentiation between participants. All ratings were skewed towards the maximum (negatively skewed) and the lowest (1) rating was only given to one employee. A scale with a range of 1 to 10, or some type of graphic scale, would be more suitable for a future study. The managers appeared to have limited experience in rating their employees and often simply used the same number for all eight KPIs in the reporting of a single participant. Some more intuitive system of rating that spreads the participants over a wider scale should be developed for future studies.

The broad range of psychological markers required for this study including three EI tests, IQ markers, BFF personality and miscellaneous scales such as motivation and optimism required a large number of short tests to be administered. In this regard bandwidth was chosen over fidelity (Cronbach & Gleser, 1965) since the administration of a full BFF personality battery such as the NEO-PIR (Costa & McCrae, 1992) and an IQ battery such as the WAIS-III (Wechsler, 1997) would have exceeded the time allocated to test the staff. Now that the tests that predict superior performance have been identified it may appropriate to include more complete scales for each of these factors. On the other hand, it may be unwise to change a working formula as alternative tests may produce different outcomes. It is probably best to retain the current tests since they have been standardised on the population of interest.

Summary

This study included three EI tests covering the spectrum of performance and self-report (trait) EI tests. The correlation between the performance and self-report tests were very low indicating that they were measuring different things; that is, they are divergent constructs. None of the three global EI scores from the tests significantly predicted the criterion, rejecting the hypothesis that EI is a unified intelligence like general intelligence. The subfactors of the MSCEIT and the EQ-i were able to significantly predict the real-world performance KPI criterion, but only by using a mix of positively and negatively weighted predictors, indicating that the EI tests were

more akin to multifactorial inventories such as the 16-PF, the NEO or the MMPI. As a final point, the provision of a positive incentive, time off from work to do the test battery appeared to significantly improved response rates and the quality of data from the participants.

Chapter 6

Overall aims and findings of this thesis

This section discusses the seven overall aims of this thesis listed in the Overview of the thesis in Chapter 1, and discusses the studies, hypotheses and results relevant to each major aim.

EI tests as predictors of a mature behavioural style. The first aim of this thesis (Chapters 1 and 2) was to determine if EI tests could identify behavioural characteristics that represented *good character* and *social maturity* as had been proposed by Goleman (1995, 1998).

In Study 1, hypothesis 1 addressed this issue and found that the only global EI score that significantly predicted the behavioural criterion was the Trait Meta-Mood Scale (TMMS); the first Mayer and Salovey scale ever published (Salovey, Mayer, Goldman, Turvey & Palfai, 1995). Another significant predictor of the behavioural criterion was the Scherer et al. Voices task; a vocal affect identification task addressing the ability to label the emotions in an audio presentation of vocal affect (Scherer, Banse & Walbott, 2001). Although this task is not typically used in EI tests, it is a veridically scored performance task that could potentially be a useful addition to an EI battery of tests. Unfortunately, the Voices task did not generalise to the DANVA Voices task utilised in Study 2. The DANVA task exhibited low reliability and did not predict the behavioural criterion; thereby raising doubts about the general predictive nature of the Voices task. Additional testing of the Scherer et al. Voices will be required to confirm the positive results obtained in Study 1. Finally in Study 1, a mix of EI tests including Voices were able to explain an additional 13% of variance over FFM personality and IQ markers. This confirmed that EI added significant incremental variance when predicting the behavioural criterion; however, only the TMMS and the Voices task were significant predictors.

In Study 2, hypothesis 2, results indicated that the TEIQue global EI score was a significant predictor of the behavioural criterion although it only explained 5% of the variance. By comparison, the single FFM factor of Extraversion explained almost three times this variance (14%) demonstrating the superior predictive validity of an 8-item Extraversion scale over the 30-item TEIQue scale for this application. By selectively choosing TEIQue and ECA subfactors through stepwise regression, the overall variance in the criterion explained rose to 32% indicating an incremental variance due to EI of 11%.

The overall conclusion drawn from these results is that EI tests do predict the behavioural criterion with significant amounts of variance above and beyond existing personality and IQ measures. However, the variance is small in comparison with the Goleman claim that "other characteristics", EI being the major one, could explain up to 80% of the variance after controlling for IQ (Goleman, 1995, p. 34). Furthermore, the practice of selecting only EI subfactors that are

predictive after results are analysed, and not according to a predefined theory, may be criticised as post-hoc science and capitalising on chance (e.g., Babyak, 2004). However, four recent studies by Bar-On used this technique to choose factors post hoc in personnel selection applications for the US Air Force and the Israeli Defence Forces, so the precedent has been set (Bar-On, Handley & Hunt, 2006). The issue of selecting only the subfactors that turn out to be predictive will be discussed further below, where the issue of whether EI is a unitary construct, or a loosely-related family of constructs is examined (e.g., Ciarrochi & Godsell, 2006).

EI as a predictor of academic performance. While IQ testing arose from, and continues to be the gold standard when it comes to predicting future academic performance (Mackintosh, 1998), the transition to university requires additional social and emotional skills for an individual to perform well in an independent learning environment (e.g., Schutte et al., 1998). Study 2, hypothesis 3, addressed this issue and found that none of the global EI scores predicted superior performance in the first-semester Psychology 104 (PSY104) results. However, after controlling for IQ and the FFM personality factors, the eight TEIQue and ECA EI subfactors explained an incremental 11% of variance. Thus, hypothesis 3 was accepted; EI does explain significant incremental variance in predicting academic performance. However, as discussed above, the predictors were chosen, post hoc, after the results were analysed because the global EI scores from TEIQue and ECA models of EI were not significant predictors.

The ability of EI test results to add significantly to predicting academic performance during the first year of university has important implications. EI testing could be included as an adjunct to the current method of using high-school results, that is, the University Admission Index (UAI) or SAT scores for admitting students. Miscellaneous subtests of Motivation and Impulse control also appeared to be significant predictors although they are not always included in EI tests such as the MSCEIT. Thus, the prediction of first-year academic performance could be significantly improved by widening the psychological domain to include EI, personality, and miscellaneous motivation questionnaires. Whether this application of EI is ever implemented depends on enrolment selection boards, who to this date have concentrated on previous high-school results and SAT-type proxy IQ test batteries.

Differences between ability EI and self-report EI. The minimal convergent validity displayed between ability EI (e.g., MSCEIT) and self-report EI tests (e.g., EQ-i) in Study 4 (Chapter 5) is too significant to continue to ignore the difference in general discussions about EI. A model of EI should be developed by researchers who have no financial gain in the outcome, to determine whether this new model of EI is a unified construct or a collection of loosely-related EI subfactors. The commercial nature, and expense of EI tests, with proprietary scoring keys that are

unavailable to the researcher, is an impediment to future research and inhibits the ability to combine the best features and factors from each test.

In Study 4 it was found that the correlations between the MSCEIT and the EQ-i were only small (r =.22, p < .05) confirming the results obtained in an earlier study (Brackett & Mayer, 2003). This represents only 5% of shared variance between the two major tests which is clearly insufficient convergent validity to claim they are both measuring EI. There have been discussions about *ability* versus *self-report* EI (e.g., Mayer, 2000b; Petrides & Furnham, 2001) but this has not clarified the issue; it has simply added to the confusion by having to preface every statement about EI with whether we are referring to ability EI or self-report EI.

In Study 4, none of the global EI scores were significant predictors of the performance KPI criterion, and the MSCEIT global score was, in fact, negatively correlated. For all the discussion about the theoretical superiority of performance versus self-report EI (e.g. Mayer et al., 2000b, 2000c; Matthews et al., 2002), this was not borne out in this study. Study 4 found that the 15 EQ-i subscales could explain 27%, almost twice the variance that the eight MCSEIT subscales could explain. In this regard, subfactors of the EQ-i self-report test were superior to the MSCEIT subfactors in predicting the criterion, and could be further enhanced with miscellaneous scales such as Motivation and Impulse control. However, there is no guidance for selecting an ability or self-report EI test at the outset of a study; the relative advantages of each can only be determined by including both tests in the study.

The burdensome overhead of having to include an ability and a self-report EI test to cover both theoretical types of EI in each study is wasteful of both the participant's and researcher's time and money. Clearly the time has arrived for both of the disparate tests to be rationalised for future applications and study. Future EI researchers would benefit from developing and sharing a range of EI items grouped into subfactors to extend EI research. In this regard Goldberg (1999) has made an excellent start by implementing the International Personality Item Pool (IPIP) that provides items for researchers. Unfortunately, the IPIP only contains 10 EI-related items; not enough for EI research and development. Overall, the lack of convergent validity between the EQ-i and the MSCEIT demonstrates that there is no theoretical support at this stage for a single EI factor as proposed by Goleman (1995).

EI, IQ and the FFM as predictors that explain all the variance. The 1995 Goleman claim that EI could potentially predict up to 80% of the variance of a performance criterion beyond IQ was based on the assumption that all other psychological factors beside IQ could be subsumed under the banner of EI (Hedlund & Sternberg, 2000; Austin & Saklofske, 2005). Goleman demonstrates a certain naivety with regard to prior research such as the MPQ (Tellegen, 1982), the MMPI (Butcher et al., 1989) the FFM of personality (Costa & McCrae, 1992), and the sixteen

personality factors (16-PF) defined by Cattell (1994) that cover many of the factors Goleman subsumed into EI (1995, 1998, 2002). In this regard, the criticism that EI is "old wine in new bottles" is justified (e.g., Salovey, 2006, p. 269).

Study 3, hypothesis 4 (Chapter 4) addressed the issue of whether there were other tests that had not been included in the EI construct that were capable of explaining additional variance over EI. It was found that *Age* and the FFM factor of *Extraversion* explained an additional 14% of variance beyond the EQ-i and TEIQue subfactors. Clearly, EI as represented by the combination of the EQ-i and TEIQue does not address all the variance in a human performance criterion. While it would be unusual to give two EI tests in a normal application of EI testing, the second EI test (TEIQue) was able to explain incremental variance beyond the EQ-i indicating that single EI tests are deficient because their subfactors are not tapping all variance available. The Age variable needs to be unravelled further to determine whether it was experience, years in the job, pragmatism that comes with maturity, or some other factor that was making it a highly significant predictor. Nevertheless, neither the EQ-i nor the TEIQue explained all the variance available, and thus had to be supplemented by additional psychological tests to maximise the variance explained.

This has implications for future EI tests as has been discussed above; it would be valuable if EI theory moved towards consolidating all the useful and predictive constructs into one test. Clearly, EI tests do not explain all the variance beyond that explained by IQ, and additional tests are required to approach the potential of explaining the 80% of remaining variance proposed by Goleman (1995).

EI as a unified construct

The claim that EI is a unified intelligence with a common general factor has been instrumental in the rapid acceptance of EI as a new construct. However, these studies showed that it is probable that EI is simply a collection of individual tests such as contained in the 16-PF, the MMPI, or the FFM of personality, that may have a socially-correct pole, but do not contain sufficient common variance to support a global factor.

The only scale that exhibited a unified construct, where the global score predicted as much variance as the separate subfactors, was the Trait Meta-Mood Scale (TMMS); a self-report scale used in Study 1 (Chapter 2). The TMMS was the first EI test published by Mayer and Salovey before they moved to ability tests with the MEIS and MSCEIT (Salovey, Mayer, Goldman, Turvey & Palfai, 1995).

In Study 2 (Chapter 3), the TEIQue and the ECA each (separately) demonstrated a unified construct in terms of confirmatory factor analysis (CFA); but not when combined. Even though the TEIQue and the ECA demonstrated reasonable convergent validity ($r = .66$, $p < .01$), compared

with the EQ-i and the MSCEIT (r = .22), this was not sufficient to confirm a single latent EI factor using CFA.

In Studies 2, 3, and 4, the subfactors of all the EI tests (TEIQue, ECA, EQ-i, and the MSCEIT) predicted significantly more variance than global EI scores. This indicates that the subfactors were acting as a loosely-related "family of constructs" rather than a "unitary construct" (see Ciarrochi & Godsell, 2006, p. 34).

Thus, apart from the TMMS there were no contemporary EI tests that demonstrated a unified construct, as proposed by Goleman, in the same way as global IQ (g) represents the underlying general intelligence factors, (c.f., Goleman, 1995). Furthermore, Bar-On appears to have tacitly admitted to the inadequacy of global EI scores in a recent study (see Bar-On, Handley & Fund, 2006). Bar-On et al. adopted the methodology of selecting only subfactors that were predictive in specific situations such as military recruiting; a tacit implication that he also finds global EI scores to have little predictive validity (Bar-On, Handley & Fund, 2006). The fact that the five models of different recruiting scenarios in his paper contained dissimilar mixes of the 15 EQ-i subfactors further demonstrates that EI as operationalised by the EQ-i is a loosely-connected set of subscales rather than a unified intelligence. In Study 4 (Chapter 5), the MSCEIT also demonstrated that it was a loosely-connected set of subscales, and not a unified intelligence. This was demonstrated by the results that showed that the individual MSCEIT subfactors explained over 30 times as much variance as the global EI score explained.

The implications of the finding that EI is a loose family of constructs is that it makes little sense reporting on an individual's total EQ or EI score, or making statements such as, "he or she is high in EI." It would be more valuable to simply interpret the subfactors that the individual has received extreme scores in. By consolidating group scores, common high and low factor scores obtained by high-performing individuals would be valuable for new recruit selection.

With regard to recruitment selection batteries, a selected set of subfactors found to be predictive in a specific environment (e.g., Bar-On, Handley & Fund, 2006) could be combined into shorter batteries than the 133-item EQ-i. This process requires a 2-step methodology: first, establish which subtests are predictive of performance in a specific environment; and second, build the battery of subtests with a scoring key to be used in the selection process. This methodology would also be applicable beyond the organizational realm, such as predicting academic performance for university applicants as discussed in Study 2.

One final point regarding the use of EI subfactors in predicting performance using regression analysis is that some subfactors were negatively related to the performance criterion. This was demonstrated in studies 2, 3 and 4 for both the EQ-i and the MSCEIT. In Study 3 it was found that the EQ-i subfactors Independence and Interpersonal relationship were negatively related

to the performance criterion with all other factors held constant. In Study 4, EQ-i subfactors Self-regard and Assertiveness were negatively related to the performance criterion. In Study 4, the MSCEIT subfactors Pictures and Changes were found to be negatively related to the criterion. This demonstrates the folly of simply summing the EI subfactors to create a global score; some factors add and some subtract from the potential variance explained by the global EI score. It also demonstrates that EI attributes are not unipolar in every situation.

It is interesting to note that the Bar-On study (Bar-On, Handley & Fund, 2006) contained no negatively-signed EI subfactors as predictors; perhaps because it would conflict with general EI theory that all subfactors are assumed to relate positively with performance. However, by using both positive and negatively signed predictors, this study was able to explain additional variance in the criterion (up to 50%); while the Bar-On et al. study using only positively-signed predictors was only able to explain approximately half this variance.

This finding, that the EI subfactors have no *a priori* positive directional relationship with the criterion, is further evidence that EI more akin to a family of subfactors rather than a unified intelligence.

EI as a superior predictor to IQ

The assertion that EI matters more than IQ to success in life and work (Goleman, 1995) was examined in Studies 3 and 4. This issue can be addressed in two ways: total EI scores versus IQ scores; or, a collection of EI subfactors selected through regression analysis versus IQ scores.

In Study 3 (Chapter 4), the total EI scores for both the EQ-i and the TEIQue demonstrated trivial and insignificant correlations with the performance criterion. Since the IQ markers also demonstrated trivial correlations with the performance criterion it was concluded that neither EI or IQ was the superior predictor. In study 4 (chapter 5), the numeric reasoning task representing one facet of IQ demonstrated a significant correlation ($r = .22$, $p < .01$) with the performance criterion, while the global EI scores (MSCEIT, EQ-i, and TEIQue) had correlations typically around 0.10, and none reached significance. The MSCEIT global EI score displayed an unexpected negative correlation of -0.13 with the performance criterion, indicting that better performers scored lower on the MSCEIT. This result clearly indicated that IQ scores were superior predictors to the MSCEIT global EI score.

Taking the alternative approach of considering EI as a collection of loosely-related subfactors, the six EI subfactors selected by stepwise regression in Study 3 (Chapter 4) explained 38% of the variance compared with IQ explaining less than 1%. In Study 4 (Chapter 5), the EQ-i subfactors explained 27% of the variance in the criterion, while IQ markers only explained 4%.

In summary, EI global factors were not superior predictors to IQ scores for the performance criterion. However, if EI subfactors are used as predictors, then EI as represented by a

loosely-related collection of subfactors is a superior predictor to IQ. However, the implication of this finding is related to the previous discussion; EI did not demonstrate the properties of a unified intelligence. Hence, it is not equivalent to compare a single marker of IQ with the total contribution of 15 EQ-i subfactors; the single, narrow marker of IQ would be expected to explain less variance than a diverse range of subfactors due to the imbalance in the number of predictors. More predictors as a rule explain more variance.

Employing EI for staff selection purposes

As an alternative to using stepwise regression analysis to choose significant predictors of the performance criterion, an independent samples *t*-test was performed on all the subfactors in the test battery. The *t*-test compared the means of the lowest and highest scorers on the performance criterion on all tests in the battery. Significant differences were found in test results for three subfactors in Study 3 (Chapter 4), and six subfactors in Study 4 (Chapter 5). The graphical representation of the results shows a smooth transition in all subfactor scores while progressing from the lowest to the highest performance rating.

The implication of this analysis is that selected EI subfactors clearly identify high and low performers. The graphical format could be readily interpreted by managers and HR personnel as it avoids some of the jargon and technical discussions surrounding regression analysis. The analysis confirmed that none of the global EI scores were able to distinguish between the high and low performers.

Alternative explanations and potential criticisms

This thesis addresses many of the issues raised by Goleman, Bar-On and Mayer et al. and provides comparative data between major EI tests using objective criteria. It is worthwhile to consider briefly criticisms that each of the major theorists may have for the findings of this study.

Goleman would probably welcome this research as it confirms the ability of EI factors to explain up to 50% of the variance in a performance criterion; that is, that EI is more important than IQ as a predictor of performance. Although he has not clearly delineated EI from previous groupings of factors such as the 16-PF and the MMPI, he would probably not agree that he had ignored them, or was naive about them. Nonetheless, by explicitly naming the EI subfactors (Goleman, 1998), and not providing any cross-reference to previous theoretical constructs there appears to be at least a tacit neglect of previous research.

Bar-On would welcome the finding that the EQ-i was the most predictive of all the EI tests of real-world performance criteria in these studies. He would probably question the inclusion of subfactors inversely related with performance on the basis that they do not meet the theoretical requirement that EQ-i subfactors have a positive directional relationship with performance criteria.

He may argue that negatively-related factors predict lower psychological well-being, and are more typically associated with failure in life. However, my response would be that superior performance in a work-related criterion does not always align with the individual's wellbeing and personal success in life; for example, the psychological strain of achieving sales targets each year, or working in the pressurised environment of a call-centre are not necessarily positively related to happiness and well-being.

Mayer et al. would be pleased that their TMMS was the most predictive test in Study 1, but would be concerned that the MSCEIT total EIQ score was negatively related to the performance criterion in Study 4. They may suggest that the performance KPI criterion was somehow problematic or ill-defined leading to this inverse relationship. My response would be that organisations set their own performance criteria and it is not for EI theorists to question their criterion; rather, EI researchers should apply their theoretical constructs to real-world situations to test predictive validity, and establish whether EI tests are a useful adjunct to current test batteries.

Overall, EI researchers and theorists may criticise the use of stepwise regression analysis to determine which EI subfactors were actually predictive of a specific criterion. Criticisms of "capitalising on chance" and "post hoc science" could be levelled at using this approach. In response, it is necessary to point out that a precedent has been set; this methodology is utilised by Bar-On et al. (2006), one of the major theorists in this area. Further, at this early stage of establishing the validity of the EI construct, no statistical method of verification or validation should be ruled out.

Issues for future research

Emotional Intelligence research and testing to date has only addressed the conscious emotional processing domain using items such as choosing the degree of emotion in faces, the correct response to a social situation (e.g., Mayer et al., 2002d), and rating one's agreement or disagreement with statements (e.g., Bar-On, 1997). However, this neglects a large body of research that addresses the dual-systems approach to emotional processing (e.g., Clore, 1994; Ellesworth, 1994; Fridja, 1994; Izard, 1994; LeDoux, 1994, 1994a; Panksepp, 1994; Scherer, 1994; Davidson & Ekman, 1994, 1994a; Zajonc, 1994; Clore & Ortony, 1994a, 2000; Damasio, 1999; Epstein, Pacini, Heier, Denes-Raj, 1996; Epstein, 1998; Bargh, 1997; Gladwell, 2005). An example of the tension created by dual-systems processing is a situation where the cerebral cortex is overridden by the primitive subcortical brain, leading to "amygdala hijacking;" a term that now enjoys widespread usage (Goleman, 1995, p. 17; LeDoux, 1994, 1994a).

Both Goleman and Mayer et al. draw a distinction between cold cognition or reasoning ability that is typically engaged in IQ testing, and hot cognitions that deal with personal and emotional matters (Goleman, 2006; Mayer, Caruso & Salovey, 1999). To obtain a complete picture

of emotion processing and EI, a necessary future task will be to determine whether there is a unified emotion processing system, or whether there are separate subsystems that interact when resolving emotional situations (e.g., LeDoux, 1994; Zajonc, 1994).

An earlier description of the separate-subsystem approach to emotional processing is found in MacLean's (1990) concept of "the triune brain;" a brain that that evolved in a layered architecture with the reptilian brain at the base, a mammalian or limbic brain surrounding the base, and the neo-cortex (bark), the primate brain, surrounding the limbic structure. The limbic brain is the centre of "hot" emotional processing; its emotionality derived from the prototypical mammalian care and concern for its offspring, while the neo-cortex represents the enabler for the ascent of man through the use of tools and language mediated by rational, logical or "cold" cognition (MacLean, 1990).

The neo-cortex is the only structure that can participate in verbal communication, although pathways to and from the limbic system ensure that it can participate after verbal information has been decoded. The evidence for three structures does not suggest that the three brains operate independently, but rather that some independent processing is possible (Maclean, 1990). However, cerebral cognition only represents the tip of the iceberg in important decisions such as choosing a mate, choosing friends, or making judgments such as in a jury or on a committee (MacLean, 1990; Damasio, 1999). In situations such as danger involving rapid detection and action, the neo-cortical conscious processing typically occurs after the fast-acting emotion-based systems have initiated a response (e.g, Damasio, 1999; Libet, 1993, 1996, 2003 ; LeDoux, 1994, 1994a; Zajonc, 1994). Although there are examples of complementary aspects of emotion and thought, such as combining rational evidence with feelings (e.g., Mayer et al., 2002c), this does not rule out independent processing within the rational versus emotional subsystems (e.g., Epstein et al., 1996, 1998).

While the inputs to the rational brain usually occur verbally or in writing, inputs to the emotional system, such as a dismissive look from a potential mate, mediated by ancestral or life experiences and the subtleties of nonverbal leakage have no such language (Clore & Ortony, 1994, 1994a, 2000; Damasio, 1999; Ekman & Davidson, 1994). The function of emotional input decoding has been compared with edge and motion-detectors in the visual system; fast, low-level cognition with important functions, but difficult to explain in linguistic terms (Griffiths, 1997, 2001). The nonverbal aspects of communicating and processing emotions appears to have been largely overlooked by key EI exponents (e.g., Bar-On, 1997), although Mayer et al. gave passing consideration to the nonverbal by including pictures and designs in the MSCEIT, and music items in their earlier EI test, the MEIS (Mayer, Salovey & Caruso, 1997).

In his latest book, Goleman reemphasizes the importance of the hot-cognition or the "low road" emotional processing system in interpersonal transactions that provide early warnings of

potential dangers (Goleman, 2006, p. 15). Although Mayer et al. acknowledge that "Emotional intelligence is a hot intelligence" (Mayer, 1999, p. 269), the test items in the MSCEIT do not reflect this; items typically only invoke responses through the cold, cortical processing system by utilising language as the medium of communication; the "high road" (Goleman, 2006). To completely investigate the emotional processing systems and to evaluate emotional responses in all their manifestations it will be necessary for EI tests of the future to include hot, or "low road" emotional intelligence items. Perhaps it is time to remind ourselves what Thorndike stated 70-years earlier; that tasks involving real people in real situations would be required to elicit the social intelligence factor (Thorndike & Stein, 1937)

Implications and concluding remarks

It would appear that EI research has only addressed a small fraction of the issues relating to measuring adaptive emotional processing that contributes to superior adaptation in work and life situations. The two dominant EI tests, the MSCEIT and the EQ-i, need to be consolidated into a single EI theoretical model; a necessary step before EI can be formally included into psychological theory as a construct worthy of the esteem that general intelligence theory, or the FFM of personality have attained. There is indeed a need to expand the list of factors that contribute to success in work and life, and in this regard EI has been a worthwhile incremental step along this path. However, the current challenge is to consolidate theories of EI, and branch out into new research areas such as measuring the wisdom and intelligence of hot emotional responses in today's environment. It is hoped that this thesis brings the current status and challenges of EI into clearer focus, and also provides some direction for future research.

Epilogue

This research began with the intention of confirming the predictive properties of global EI scores; especially since Goleman indicated that EI could explain up to 80% of the variance of success in work and life criteria. The four studies conducted in this research, using widely-recognised EI tests, found that global EI scores explained insignificant quantities of variance; typically less than 5%. The four studies found no evidence to support the Goleman assertion that a single, global EI factor was significantly related to work and life success. The studies confirmed that EI was dissimilar to general intelligence; global EI scores could not adequately represent an individual's maturity and success at work.

However, this research did find that EI subfactors such as Self-control, Independence, and BFF personality factors such as Extraversion could explain up to 50% of the variance in a success criterion; thus approaching the Goleman claim of explaining up to 80% of the variance, and certainly surpassing the variance that IQ can explain. But, some EI subfactors were negatively

related to superior performance thus refuting the claim that EI subfactors are unidirectionally related to success and performance. EI subfactors appear to behave in a similar manner to other multifactorial inventories such as the MPQ, MMPI, 16-PF and BFF personality theory. Thus, EI batteries can explain additional variance in a performance criterion by using EI subfactors as a multifactorial inventory, but the concept of a single, global EI score was found to be without merit.

Summary

Until 1995, IQ was considered the best predictor of work performance for applicants with no experience. Further, Herrnstein and Murray (1994) argued that IQ was immutable and that genetic inheritance largely determined one's station in life, and success and satisfaction in a variety of realms. However, there are numerous examples of individuals with poor scholastic records, and possibly lower than average IQ scores, who have been successful in business and in life.

Against the deterministic background of IQ, Goleman's book (1995) *Emotional Intelligence: Why it can matter more than IQ*, captured the attention of a wide audience, especially in the industrial-organizational and academic sectors. Twelve years on, it is not uncommon to hear of organizations evaluating employees on their Emotional Intelligence (EI) skills. However, there is no common model for EI, and the two leading instruments, the EQ-i and the MSCEIT, do not measure the same thing; an issue of construct validity. Many publications describing the virtues of EI select their own models of EI, thereby producing results and outcomes that cannot be generalised to any other industry or application.

A further issue with EI is whether it can be represented by a single score, like IQ, or whether it comprises a family of loosely-related abilities and skills, the scores of which cannot be meaningfully summed to produce a total EI score. Overall, there is a powerful lobby of EI practitioners working with various models of EI, using different tests, none of which conform to a single, validated model of EI.

This thesis attempts to address the wide disparity of definitions and tests of EI by providing some concrete grounding of total EI and subfactor scores with behavioural, academic, and corporate success criteria. The thesis consists of four studies; two of which compare EI scores with behavioural criteria of "maturity" and "good character" as proposed by Goleman (1995). The remaining studies examine EI as a predictor of academic success, and success at work; specifically, field sales representatives, and customer service representatives in a call-centre environment. All studies follow a similar methodology involving the administration of EI, personality, IQ, and miscellaneous emotion-related tests and comparing the results with a behavioural, academic, or a work-related criterion.

The overall finding of this thesis is that EI in its various guises can be usefully employed to predict social maturity, academic success, and success in the workplace. However, total or global EI scores were not significant predictors of success; in fact, in some cases they were inversely related to success. Thus, EI as measured by the two leading EI tests, the MSCEIT and the EQ-i, did not demonstrate the properties of a unified construct; global EI scores were inferior to EI subfactors when it came to explaining variance in the criterion.

This thesis concludes that EI does not exhibit the properties of a unitary construct; it manifests as a family of loosely-related abilities that can, however, be successfully employed in predicting superior performance. However, the EI subfactors that predicted superior performance were often found to be negatively related to the performance criterion.

The overall conclusion of the thesis is that total EI scores and subfactors relating to superior performance in a specific application cannot be determined *a priori*. Determining the EI factors that predict superior performance requires a prior study of the existing population to determine the relevant EI subfactors, and their relative weighting and directionality, applicable to the criterion of success chosen for that industry or application.

References

ABS. (2001). Australian Bureau of Statistics 2001 *Census of population and housing (Sydney)*, from www.abs.gov.au

Allport, G. W. (1958). What units shall we deploy. In G. Lindzey (Ed.), *Assessment of human motives*. New York: Holt, Rinehart and Winston.

Allport, G. W., & Odbert, H. S. (1936). Trait-names. A psycho-lexical study. *Psychological Monographs, 47*(211), 171.

Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Upper Saddle River, NJ: Prentice Hall.

APA. (2006). PsycINFO: American Psychological Association: http://www.apa.org/psycinfo/.

Arbuckle, J. L. (2005). Amos 6.0 user's guide. Chicago: SPSS.

Arbuckle, J. L. (2005a). Amos (Version 6.0) [Computer Program]. Chicago: SPSS.

Austin, E. J., & Saklofske, D. H. (2005). Far too many intelligences? On the communalities and differences between social, practical, and emotional intelligences In R. Schulze & R. D. Roberts (Eds.), *Emotional Intelligence: An international handbook*. Washington: Hogrefe & Huber.

Babyak, M. A. (2004). What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine, 66*, 411-421.

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty item Toronto Alexithymia Scale-I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38*(1), 23-32.

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994a). The twenty item Toronto Alexithymia Scale-II. Convergent, discriminant and concurrent validity. *Journal of Psychosomatic Research, 38*(1), 33-40.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer (Ed.), *The automaticity of everyday life* (pp. 1-62). Mahwah, NJ: Lawrence Erlbaum Associates.

Bar-On, R. (1997). *BarOn Emotional Quotient Inventory: Technical manual*. North Tonawanda, NY: Multi-Health Systems.

Bar-On, R. (1997a). *BarOn Emotional Quotient Inventory - Item booklet (133)*. New York: Multi-Health Systems (MHS).

Bar-On, R. (2000). Emotional and social intelligence: Insights from the Emotional Quotient Inventory. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence*. San Francisco: Jossey-Bass.

Bar-On, R. (2006a). The Bar-On model of emotional-social intelligence (ESI). *Psicothema, 18*(supl.), 13-25.

Bar-On, R., Handley, R., & Fund, S. (2006). The impact of Emotional Intelligence on performance. In V. U. Druskat, F. Sala & G. Mount (Eds.), *Linking Emotional Intelligence and performance at work*. Marwah, NJ: Lawrence Erlbaum Associates.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hove, UK: Lawrence Erlbaum.

Boring, E. G. (1923). Intelligence as the tests test it. *New Republic, 35*, 35-37.

Boyatzis, R. E., Goleman, D., & Rhee, K. S. (2000). Clustering competence in Emotional Intelligence: Insights from the Emotional Competence Inventory. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence*. San Francisco: Jossey-Bass.

Boyatzis, R. E., Goleman, D., & Rhee, K. S. (2006). Emotional Competence Inventory (ECI): Technical manual. Retrieved Nov 23, 2006, 2006, from http://www.eiconsortium.org/research/ECI_2.0_Technical_Manual_v2.pdf

Brackett, M. A., & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin, 29* (9), 1147-1158.

Brackett, M. A., Rivers, S. E., Shiffman, S., Lerner, N., & Salovey, P. (2006). Relating emotional abilities to social functioning: A comparison of self-report and performance measures of emotional intelligence. *Journal of Personality and Social Psychology, 91*(4), 780-785.

Bradberry, T., & Greaves, J. (2004). Emotional Intelligence Appraisal - Technical manual. from http://www.eiconsortium.org/research/EIA_Technical_Manual.pdf

Braithwaite, E. R., & Clavell, J. (Writer) (1967). *To Sir, with love*. Columbia British Productions.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory - 2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

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

Carroll, L. (2004). *Alice in wonderland*. NY: Gramercy Books.

Caruso, D., & Wolfe, C. J. (2001). Emotional Intelligence in the workplace. In J. Ciarrochi, J. P. Forgas & J. D. Mayer (Eds.), *Emotional intelligence in everyday life: A scientific enquiry* (pp. 150-167). Philadelphia, PA, USA: Psychology Press: Taylor & Francis.

Caruso, D. R., Salovey, P., & Mayer, J. D. (2003). Emotional intelligence and emotional leadership. In R. E. Riggio & S. E. Murphy (Eds.), *Multiple intelligences and leadership* (pp.55-73). Mahwah, NJ: Erlbaum.

Cattell, R. B., Cattell, K. A., & Cattell, H. E. P. (1994). *The Sixteen Personality Factor Questionnaire (16PF) 5th ed*: USA: Institute for Personality and Ability Testing.

Cherniss, C. (2000). Social and Emotional Competence in the workplace. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of Emotional Intelligence*. San Francisco: Jossey-Bass.

Cherniss, C. (2001). Emotional Intelligence and organizational effectiveness. In C. Cherniss & D. Goleman (Eds.), *The emotionally intelligent workplace* (pp.3-12). San Francisco: Jossey-Bass.

Cherniss, C. (2004). The business case for emotional intelligence [Electronic Version]. Retrieved 11-Jan-2007 from www.eiconsortium.org.

Cherniss, C., & Goleman, D. (2001a). *The emotionally intelligent workplace*. San Francisco, CA: Jossey-Bass.

Child, D. (1976). *The essentials of Factor Analysis*. London: Holt, Reinhart & Winston.

Child, D. (2006). *The essentials of Factor Analysis*. London: Continuum International Publishing Group; 3rd edition

Christie, R., & Geis, F. (1970). Scale construction. In *Studies in machiavellianism* (pp. 10-34). NY: Academic Press.

Cialdini, R. B. (2001). *Influence: Science and practice*. Boston, MA: Allyn and Bacon.

Ciarrochi, J., Forgas, J., & Mayer, J. D. (2001). *Emotional Intelligence in everyday life: A scientific enquiry*. Philadelphia, PA: Psychology Press.

Ciarrochi, J., & Godsell, C. (2006). Mindfulness-based emotional intelligence: Research and training. In V. U. Druskat, F. Sala & G. Mount (Eds.), *Linking Emotional Intelligence and performance at work* (pp. 21-52). Marwah, NJ: Lawrence Erlbaum Associates.

Clore, G. L. (1994). Why emotions require cognition: Fundamental questions. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp.181-191). New York: Oxford University Press.

Clore, G. L. (1994a). Why emotions are never unconscious. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 285-290). New York: Oxford University Press.

Clore, G. L., & Ortony, A. (2000). Cognition in emotion: Always, sometimes, or never? In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 24-61). New York, NY: Oxford University Press.

Cohen, J. (1988). *Statistical power for the behavioural sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.

Collins, V. L. (2002). EI and leadership success. *Dissertation Abstracts International, 62*(11-B).

Costa, P. T., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PIR) and NEO Five-Factor Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.

Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions*. Urbana, Il: University of Illinois press.

Damasio, A. (1999). *The feeling of what happens*. New York: Harcourt, Inc.

Davidson, R. J., & Ekman, P. (1994). Afterword: Can emotions be nonconscious? In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 298-300). New York: Oxford University Press.

Davidson, R. J., & Ekman, P. (1994a). Afterword: What are the minimal cognitive prerequisites for emotion? In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 232-234). New York: Oxford University Press.

Davies, M., Stankov, L., & Roberts, R. D. (1998). Emotional Intelligence: In search of an elusive construct. *Journal of Personality and Social Psychology, 75*(4), 989-1015.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalogue of Selected Documents in Psychology, 10*(85).

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multi-dimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113-126.

Davis, M. H., & Kraus, L. A. (1991). Dispositional empathy and social relationships. In W. H. Jones & D. Perlman (Eds.), *Advances in personal relationships*. Greenwich, CN: JAI Press.

Dawda, D., & Hart, S. D. (2000). Assessing emotional intelligence: Reliability and validity of the Bar-On Emotional Quotient Inventory (EQ-i) in university students. *Personality and Individual Differences, 28*, 797-812.

Deeter-Schmelz, D. R., & Sojka, J. Z. (2003). Developing effective salespeople: Exploring the link between Emotional Intelligence and sales performance. *The International Journal of Organizational Analysis, 11*(3), 211-220.

Dekker, A. (2005). Conceptual distance in social network analysis [Electronic Version]. *Journal of Social Structure*, 6. Retrieved 20/1/2007 from http://www.cmu.edu/joss/content/articles/volume6/dekker/.

Dulewicz, V., & Higgs, M. (2000). Emotional intelligence: A review and evaluation study. *Journal of Managerial Psychology, 15*(4), 341-372.

Ekman, P. (2003). Micro expression training tool (METT). from www.emotionsrevealed.com

Ekman, P. (2003a). *Emotions Revealed*. London: The Orion Publishing Group.

Ekman, P., & Davidson, R. (1994). *The nature of emotion: Fundamental questions*. New York: Oxford University Press.

Ekman, P., & Friesen, W. V. (1978). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ellseworth, P. C. (1994). Levels of thought and levels of emotion. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 192-196). New York: Oxford University Press.

Emmerling, R., & Goleman, D. (2006). The EI Consortium for research on emotional intelligence in organizations. from http://eiconsortium.org

Epstein, S. (1998). *Constructive thinking: The key to Emotional Intelligence*. Westport, CT: Praeger Publishers.

Epstein, S., Pacini, R., Heier, H., & Denes-Raj, V. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology, 71*(2), 390-405.

Eysenck, H. J. (1959). Manual of the Maudsley Personality Inventory (MPI) London: University of London Press.

Fagan, P. F. (2006). Why religion matters even more: The impact of religious practice on social stability [Electronic Version]. Retrieved 2/4/2007 from http://www.heritage.org/Research/Religion/bg1992.cfm.

Field, A. (2005). *Discovering statistics using SPSS (2nd Edition)*. London: Sage Publications Ltd.

Fletcher, G. J. O., & Fitness, J. (1990). Occurrent social cognition in close relationship interaction: The role of proximal and distal variables. *Journal of Personality and Social Psychology, 59*(3), 464-474.

Ford, M. E., & Tisak, M. S. (1983). A further search for Social Intelligence. *Journal of Educational Psychology, 75*(2), 196-206.

Forgas, J., Mayer, J. D., & Ciarrochi, J. (2001). *Emotional Intelligence in everyday Life*: Psychology Press.

Frijda, N. H. (1994). Emotions require cognitions, even if simple ones. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 197-202). New York: Oxford University Press.

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioural consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology, 60*(5), 773-794.

Funder, D. C., & Sneed, C. D. (1993). Behavioural manifestations of personality: An ecological approach to judgemental accuracy. *Journal of Personality and Social Psychology, 64*(3), 479-490.

Gardner, H. (1983). *Frames of mind*. New York: Basic books.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York: Basic books.

Gibbs, N. (1995, October 2). The EQ Factor. *Time*.

Gladwell, M. (2005). *Blink: The power of thinking without thinking*. London: Penguin.

Goetz, T., Frenzel, A. C., Pekrun, R., & Hall, N. (2005). Emotional learning in the context of learning and achievement. In R. Schulze & R. D. Roberts (Eds.), *Emotional Intelligence: An international handbook*. Washington: Hogrefe & Huber.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Journal of Personality and Social Psychology, 59*, 1216-1229.

Goldberg, L. R. (1994). Mini-markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment, 63*(3), 506-516.

Goldberg, L. R. (1999). The international personality item pool: A scientific collaboratory for the development of advanced measures of personality traits and other individual differences (http://ipip.ori.org/).

Goleman, D. (1995). *Emotional Intelligence: Why it can matter more than IQ*. London: Bloomsbury.

Goleman, D. (1996). Emotional intelligence: a new vision for educators (Videorecording). Port Chester, N.Y. : National Professional Resources, Inc.

Goleman, D. (1998). *Working with Emotional Intelligence*. London: Bloomsbury.

Goleman, D. (2001a). Emotional Intelligence: Issues in paradigm building. In C. Cherniss & D. Goleman (Eds.), *The emotionally intelligent workplace* (pp. 13-26). San Francisco: Jossey-Bass.

Goleman, D. (2001b). An EI based theory of performance. In C. Cherniss & D. Goleman (Eds.), *The emotionally intelligent workplace* (pp. 27-44). San Francisco: Jossey-Bass.

Goleman, D. (2003). Emotional intelligence appraisal. San Diego: TalentSmart Service.

Goleman, D. (2006). *Social intelligence: The new science of human relationships* NY: Bantam Books.

Goleman, D., Boyatzis, R., & McKee, A. (2002). *Primal leadership: Realizing the power of emotional intelligence*. Boston, MA: Harvard Business School Press.

Griffiths, P. E. (1997). *What emotions really are*. Chicago: University of Chicago Press.

Griffiths, P. E. (2001). *Basic emotions, complex emotions, Machiavellian emotions*. Department of History and Philosophy of Science (pauleg@pitt.edu): University of Pittsburgh.

Guttman, L., & Levy, S. (1991). Two structural laws for intelligence tests. *Intelligence, 15*, 79-103.

Harman, H. A. (1976). *Modern factor analysis*. Chicago: University of Chicago Press.

Hathaway, S. R., McKinley, J. C., Butcher, J. N., Dahlstrom, W. G., Graham, J. R., & Tellegen, A. (1989). *Minnesota Multiphasic Personality Inventory-2*. Minneapolis: The University of Minnesota Press.

Hedlund, J., & Sternberg, R. J. (2000). Too many intelligences? Integrating Social, Emotional, and Practical Intelligence. In R. Bar-On & J. D. A. Parker (Eds.), *The Handbook of Emotional Intelligence* (pp.136-167). San Francisco: Jossey-Bass.

Hein, S. (2006). Emotions. from http://eqi.org/

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: The Free Press.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33*(3), 307-317.

Howell, D. C. (1997). *Statistical methods for psychology (4th. ed.)*. Belmont, CA.: Duxbury Press.

Hunt, R. H. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology, 130*(4), 658-680.

Isenberg, N., Silbersweig, D., Engelein, A., Emmerich, S., Malavade, K., Beattie, B., et al. (1999). Linguistic threat activates the human amygdala. *Proceedings of the National Academy of Science, USA, 96*(August), 10456-10459.

Izard, C. E. (1994). Cognition is one of four types of emotion activating systems. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 203-207). New York: Oxford University Press.

Izard, C. E. (2001). Emotional intelligence or adaptive emotions. *Emotion, 1*(3), 249-257.

Jackson, C. J. (2001). Comparison between Eysenck's and Gray's models of personality in the prediction of motivational work criteria. *Personality and Individual Differences, 31*, 129-144.

Johnson, J. A., Cheek, J. M., & Smither, R. (1983). The structure of empathy. *Journal of Personality and Social Psychology, 45*(6), 1299-1312.

Kihlstrom, J. F., & Kantor, N. (2000). Social intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 359-379). New York: Cambridge University Press.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: The University of Chicago press.

LeDoux, J. E. (1994). Cognitive-emotional interactions in the brain. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp.216-223). New York: Oxford University Press.

LeDoux, J. E. (1994a). Emotional processing, but not emotions, can occur unconsciously. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 291-292). New York: Oxford University Press.

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

Legree, P. J., Martin, D. E., & Psotka, J. (2000). Measuring cognitive aptitude using unobtrusive knowledge tests: A new survey technology. *Intelligence, 28*(4), 291-308.

Libet, B. (1993). *Neuropsyiology of consciousness*. Boston, MA: Birkhauser.

Libet, B. (1996). Neural time factors and unconscious mental functions. In S. R. Hameroff, A. W. Kaszniak & A. C. Scott (Eds.), *Towards a science of consciousness: The first Tucson discussions and debates*. Cambridge, MA: The MIT Press.

Libet, B. (2003). Timing of conscious experience. *Consciousness and cognition, 12*, 321-331.

Mackintosh, N. J. (1998). *IQ and human intelligence*. New York: Oxford University Press.

MacLean, P. D. (1990). *The triune brain in evolution*. New York: Plenum Press.

Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., et al. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behavior, 24*(3), 179-209.

Matthews, G., Roberts, R. D., & Zeidner, M. (2004). Seven myths about Emotional Intelligence. *Psychological Inquiry, 15*(3), 179–196.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional Intelligence: Science and myth*. Cambridge, MA: MIT Press.

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 27*(4), 267-298.

Mayer, J. D., Caruso, D. R., & Salovey, P. (2000). Selecting a measure of emotional intelligence: The case for ability scales. In R. Bar-On & J. D. A. Parker (Eds.), *The Handbook of Emotional Intelligence*. San Francisco: Jossey-Bass.

Mayer, J. D., & Cobb, C. D. (2000a). Educational policy on emotional intelligence: Does it make sense? *Educational Policy Review, 12*(2), 163-183.

Mayer, J. D., & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence, 17*, 433-442.

Mayer, J. D., Salovey, P., & Caruso, D. R. (1997). *Multifactor Emotional Intelligence Scale (MEIS)*. Needham, MA: Virtual Knowledge (CD-ROM).

Mayer, J. D., Salovey, P., & Caruso, D. R. (2000b). Models of Emotional Intelligence. In R. J. Sternberg (Ed.), *The handbook of intelligence* (pp. 396-420). New York: Cambridge University Press.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2002c). *Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) - User's manual*. Toronto, Canada: Multi Health Systems.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2002d). *Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) - Item booklet*. Toronto, Canada: Multi Health Systems.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2001). Emotional Intelligence as a standard intelligence. *Emotion, 1*(3), 232-242.

Mayer, J. D., & Stevens, A. A. (1994). An emerging understanding of the reflective (meta-) experience of mood. *Journal of Research in Personality, 28*, 351-373.

Mazurski, E. J., & Bond, N. W. (1993). A new series of slides depicting facial expressions of affect: A comparison with pictures of facial affect series. *Australian Journal of Psychology, 45*(1), 41-47.

McCrae, R. R. (2000). Emotional Intelligence from the perspective of the Five Factor model of personality. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of Emotional Intelligence* (pp. 263-276). San Francisco: Jossey-Bass.

McCrae, R. R., & John, O. P. (1992). An introduction to the Five Factor model and its applications. *Journal of Personality, 60*(2), 175-215.

MHS. (2006). *Multi-Health Systems Inc.* North Tonawanda, NY, http://www.mhs.com/.

Michell, J. (1990). *An introduction to the logic of psychological measurement*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Mill, J. S. (2002). *A system of logic (1843)*. Honolulu, Hawaii: University Press of the Pacific

Mischel, W., Ebbesen, E. B., & Zeiss, A. R. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology, 21*(2), 204-218.

Murphy, K. R., & Davidshofer, C. O. (1998). *Psychological testing: Principles and applications (Fourth Edition)*. New Jersey: Prentice Hall.

Myers, I. B. (1980). *Myers-Briggs Type Inventory*. Consulting Psychologists Press.

Newsome, S., Day, A., & Catano, V. (2000). Assessing the predictive validity of emotional intelligence. *Personality and Individual Differences, 29*(6), 1005-1016.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*, 1-32.

Nowicki, S. (1998). *Diagnostic analysis of nonverbal accuracy (DANVA2)*: Dyssemia Inc. (www.dyssemia.com).

Nowicki, S., Jr., & Duke, M. P. (2001). *Nonverbal receptivity: The diagnostic analysis of nonverbal accuracy (DANVA)*. Mahwah, NJ: Erlbaum.

Nowicki, S., & Baum. (1998). *The DANVA2 adult paralanguage test (DANVA2-AP)*. Dyssemia Inc. (www.dyssemia.com).

O'Connor, R., & Little, I. (2003). Revisiting the predictive validity of emotional intelligence: Self report versus ability based measures. *Personality and Individual Differences, 35*(8), 1893-1902.

Onwuegbuzie, A. J., & Daniel, L. G. (2003). Typology of analytical and interpretational errors in quantitative and qualitative educational research. *Current Issues in Education, 6*(2).

Oswald, W. D. (1987). *Zahlen-Verbindungs-Test ZVT*. Göttingen: Hogrefe.

Palmer, B., & Stough, C. (2001). *Swinburn University Emotional Intelligence test: Interim technical manual - Version 2*. Melbourne: Swinburne University.

Panitz, E. (1989). Psychometric investigation of the MACH IV scale measuring Machiavellianism. *Psychological Reports, 64*, 963-968.

Panksepp, J. (1994). A proper distinction between affective and cognitive process is essential for neuroscientific progress. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 224-226). New York: Oxford University Press.

Payne, W. L. (1986). A study of emotion: developing emotional intelligence (Doctoral Dissertation) [Electronic Version] from http://eqi.org/payne.htm.

Petrides, K. V., Frederickson, N., & Furnham, A. (2004). The role of trait emotional intelligence in academic performance and deviant behaviour at school. *Personality and Individual Differences, 36*, 277-293.

Petrides, K. V., & Furnham, A. (2000). Gender differences in measured and self-estimated Trait Emotional Intelligence. *Sex Roles, 42*(5/6).

Petrides, K. V., & Furnham, A. (2000a). On the dimensional structure of emotional intelligence. *Personality and Individual Differences, 29*, 313-320.

Petrides, K. V., & Furnham, A. (2001). Trait Emotional Intelligence: Psychometric investigation with reference to established trait taxonomies. *European Journal of Personality, 15*, 425-448.

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

Petrides, K. V., & Furnham, A. (2004). *Technical manual of the Trait Emotional Intelligence Questionnaire (TEIQue)*. London: University of London, Institute of Education.

Petrides, K. V., & Furnham, A. (2006). *The role of trait emotional intelligence in a gender-specific model of organizational variables*. Symposium conducted at the 11th biennial meeting of the International Society for the Study of Individual Differences, Graz, Austria.

Petrides, K. V., Perez, J. C., & Furnham, A. (2003). *The Trait Emotional Intelligence Questionnaire (TEIQue). A measure of emotional self-efficacy.* Paper presented at the Symposium

conducted at the 11th biennial meeting of the International Society for the Study of Individual Differences, Graz, Austria.

Petrides, K. V., Perez-Gonzalez, J. C., & Furnham, A. (2007). On the criterion and incremental validity of trait emotional intelligence. *Cognition and Emotion, 21*(1), 26-55.

Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge*. London: Routledge.

Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence*. London: H. K. Lewis.

Reber, A. S. (1993). *Implicit learning and tacit knowledge*. New York: Oxford University Press.

Reid, J. R. (2002). *Performance-based emotional intelligence: Identifying measurement problems and suggested remedies using scaling theories.* Unpublished Honours thesis, University of Sydney, NSW, Australia.

Reid, J. R. (2004, 3rd December, 2004). *Criteria for emotional intelligence test development.* Paper presented at the 3rd Australian conference for personality and individual differences (ACPID), University of Ballarat, Victoria.

Reid, J. R. (2005, 18th July, 2005). *Using a criterion to provide a foundation for EI test development.* Paper presented at the The 12th biennial meeting of the international society for the study of individual differences (ISSID), University of Adelaide, South Australia.

Reid, J. R. (2006). Emotional competence assessment (ECA) scale: (unpublished test battery) Macquarie University.

Reitan, R. M. (1985). *The Halstead-Reitan neuropsychological test battery : Theory and clinical interpretation.* Tucson, Arizona: Neuropsychology Press.

Roberts, R. D. (1998). *Working memory, attention, speed of procession, and intelligence (WASP-I).* Brooks, TX: Armstrong Laboratory.

Roberts, R. D. (1999). *Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism Index Condensed (OCEANIC): Preliminary technical report draft 2*. Sydney: Sydney University.

Roberts, R. D. (2000). *Openness Conscientiousness Extraversion Agreeableness Neuroticism Index Condensed (OCEANIC) Test Manual*. Sydney, Australia: Entelligent Testing Products.

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*(4), 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*(3), 196-231.

Rozell, E. J., Pettijohn, C. E., & Parker, R. S. (2004). Customer-oriented selling: Exploring the roles of Emotional Intelligence and organizational commitment. *Psychology and Marketing, 21*(6), 405-424.

Rushton, J. P. (1997). *Race, evolution, and behavior*. New Brunswick, NJ: Transaction Publishers.

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.

Saklofske, D. H., & Parker, J. (2005). Emotional intelligence and education. In C. Stough, D. H. Saklofske & K. Hansen (Eds.), *Emotional intelligence international symposium 2005*. Swinburne University: Tertiary press.

Salovey, P. (2006). The agenda for future research. In V. U. Druskat, F. Sala & G. Mount (Eds.), *Linking Emotional Intelligence and performance at work*. Marwah, NJ: Lawrence Erlbaum Associates.

Salovey, P., Brackett, M. A., & Mayer, J. D. (2004). *Emotional intelligence: Key readings on the Mayer and Salovey model*. Port Chester, NY, US: Dude Publishing

Salovey, P., & Mayer, J. D. (1990). Emotional Intelligence. *Imagination, Cognition and Personality, 9*(3), 185-211.

Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., & Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale. In J. W. Pennebaker (Ed.), *Emotion disclosure and health*. Washington, DC: American Psychological Association.

Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment, 63*(3), 506-516.

Scherer, K. R. (1994). An emotion's occurrence depends on the relevance of an event to the organism's goal/need hierarchy. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 227-231). New York: Oxford University Press.

Scherer, K. R. (2002). Vocal emotion recognition task. Geneva. Supplied by: Tanja.Banziger@pse.unige.ch

Scherer, K. R., Banse, R., & Walbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76-92.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262-274.

Schulze, R., & Roberts, R. D. (2005). *Emotional Intelligence: An international handbook*. Washington: Hogrefe & Huber.

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

Seligman, M. E. P. (1990). *Learned optimism*. Sydney, Australia: Random House.

Sharma, D., & McKenna, F. P. (2001). The role of time pressure on the emotional Stroop task. *British Journal of Psychology, 92*, 471-481.

Spearman, C. (1904). "General Intelligence" objectively determined and measured. *American Journal of Psychology, 15*, 201-293.

Spencer, L. M. (2001). The economic value of EI competencies and EIC-based HR programs. In C. Cherniss & D. Goleman (Eds.), *The emotionally intelligent workplace*. San Francisco: Jossey-Bass.

Stadler, M. A., & Frensch, P. A. (1998). *Handbook of implicit learning*. Thousand Oaks, CA: Sage Publications.

Stankov, L. (1997). *The Gf/Gc Quickie test battery*: (unpublished test battery) The University of Sydney - School of Psychology.

Stankov, L., & Roberts, R. D. (2001). Gf/Gc battery builder. Sydney University: (unpublished test battery) Sydney University - Department of Psychology.

Sternberg, R. J. (2000). *Handbook of intelligence*. Cambridge: CUP.

Sternberg, R. J. (2000a). *Practical intelligence in everyday life*. Cambridge: CUP.

Sternberg, R. J., & Hedlund, J. (2000). Too many intelligences? Integrating social, emotional, and practical intelligence. In R. E. P. Bar-On, James D. A. (Ed). (Ed.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (pp. 136-167). New Haven, CT: Yale University.

Strube, M. J. (2002). Reliability and generalizability theory. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics*. Washington, DC: APA.

Tellegen, A. (1982). *Multidimensional Personality Questionnaire (MPQ)*. Minnesota: University of Minnesota, Dept of Psychology.

Tellegen, A., Patrick, & Curtin. (2002). Development of a brief MPQ. *Psychological Assessment, 14*(2), 150-163.

Thorndike, E. L. (1920). Intelligence and its uses. *Harper's Magazine*(140), 227-235.

Thorndike, R. L., & Stein, S. (1937). An evaluation of the attempts to measure social intelligence. *Psychological Bulletin, 34*(5), 275-285.

Vernon, P. A. (1993). Der Zahlen-Verbindungs-Test and other trail-making correlates of general intelligence. *Personality and Individual Differences, 14*(1), 35-40.

Vickers, D., & Lee, M. D. (1998). Never cross the path of a traveling salesman: The neural network generation of Halstead-Reitan trail making tests. *Behaviour Research Methods, Instruments and Computers, 30*(3), 423-431.

Watkin, C. (2000). Developing Emotional Intelligence. *International Journal of Selection and Assessment, 8*(13), 89-92.

Wechsler, D. (1939). *The measurement and appraisal of adult intelligence*. Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1940). Non-intellective factors in general intelligence. *Psychological Bulletin, 37*, 444-445.

Wechsler, D. (1950). Cognitive, conative and non-intellective intelligence. *American Psychologist 5*, 78-81.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*. Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Wilson, D. S., Near, D., & Miller, R. R. (1996). Machiavellianism: A synthesis of evolutionary and psychological literatures. *Psychological Bulletin, 119*(2), 285-299.

Zajonc, R. B. (1994). Evidence for nonconscious emotions. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 293-297). New York: Oxford University Press.

Zook, A., & Sipps, G. J. (1985). Reliability data and sex differences with a gender-free Mach IV. *The Journal of Social Psychology, 126*(1), 131-132.