The Equivalence of Online and Paper-Pencil Measures of Emotional Intelligence

Barbara B. Meyer, Susan E. Cashin and William V. Massey
University of Wisconsin, Milwaukee
USA

1. Introduction

Improved access to the Internet along with technological advances in hardware and software are prompting social scientists to move their research pursuits from the laboratory into cyberspace. While evidence suggests that online research is equivalent if not superior to traditional offline (i.e., paper-pencil [PP]) methods (Buchanan & Smith, 1999; Campos et al., 2011; Lonsdale et al., 2006; Meyerson & Tryon, 2003; Naus et al., 2009; Preckel & Thiemann, 2003), caution is urged in summarily generalizing research methods and results from one format to another. To address one of the key limitations of Internet research (i.e., measurement error), scholars have begun to examine the reliability and validity of online test formats. In studies of matched or paired samples who completed electronic and PP measures of various psychological constructs (e.g., personality, burnout, intellectual giftedness), results suggest comparable psychometric properties, factor structure, and outcomes across the data collection methods (Buchanan & Smith, 1999; Campos et al., 2011; Lonsdale et al., 2006; Meyerson & Tryon, 2003; Naus et al., 2009; Preckel & Thiemann, 2003).

While the literature reviewed above supports the general reliability and validity of Internet data collection in psychological research, it cannot be assumed that all psychological tests are equivalent across methods. Before we can consider online data collection as standard operating procedure and confidently combine the results of data collected from mixed formats, the influence of survey format on responses to specific assessment inventories must be evaluated. Little is known, for example, about the online assessment of the popular psychological construct emotional intelligence (EI).

2. Emotional intelligence

Since its introduction (Salovey & Mayer, 1990), EI, or the ability to perceive, utilize, understand, and manage emotions, has received consistent attention in both the scholarly and popular press. Despite ongoing debate over theoretical underpinnings and assessment techniques, scholars and practitioners alike have remained steadfast in their quest to use EI to describe, predict, and explain various outcomes (e.g., depression and anxiety, problem behavior, substance use, employee and customer satisfaction, etc.) across domains (Kafetsios & Zampetakis, 2008; Martins et al., 2010; Siu, 2009; Trinidad & Johnson, 2002). Lack of agreement among applied researchers as to the conceptualization and assessment of
EI can result in, among other things, different EI profiles or recommendations for the same person (Brackett & Mayer, 2003). In an effort to better understand the need for a consistent approach to EI assessment in the applied context, we review below the theoretical models and corresponding assessment inventories typically used in EI research.

2.1 Conceptualization and assessment of emotional intelligence

Most of the research on the application of EI is informed by one of two models: the mixed model or the ability model. Mixed models suggest that EI encompasses both mental abilities (i.e., emotional self-awareness, empathy, problem-solving, impulse control) and self-reported personality characteristics (i.e., mood, genuineness, warmth) (Sternberg et al., 2000). Conversely, the ability model of EI represents an intelligence involving emotion, and the ability to use the information encoded in emotion to direct cognition and motivate behavior. The most commonly utilized mixed models and ability approaches, along with their associated assessment tools, are reviewed below.

2.1.1 Goleman’s mixed model of emotional intelligence

According to Goleman (1995), EI includes “abilities such as being able to motivate oneself and persist in the face of frustrations; to control impulse and delay gratification; to regulate one’s moods and keep distress from swamping the ability to think; to empathize and to hope” (Goleman, 1995, p.34). These capabilities are present in 20 competencies that fall within four separate domains: self-awareness, self-management, social awareness, and relationship management. Although Goleman’s work is popular among laypersons and scholars who are new to the study of EI, comprehensive evaluation is made difficult because to date, little if any peer-reviewed research has been informed by Goleman’s model or related measurement tool (Landy, 2005; Matthews et al., 2004).

The aforementioned assessment tool, the Emotional Competence Inventory (ECI) – Version 2, is a 110-item self-report measure of 20 behavioral competencies within four domains. Both the validity and reliability of the instrument have been called into question (Conte, 2005; Matthews et al., 2004), including the identification of considerable overlap between the ECI and measures of the Big Five personality factors (i.e., neuroticism, extraversion, openness, agreeableness, conscientiousness) (Conte, 2005; Matthews et al., 2004; Van Rooy & Viswesvaran, 2004). Since few independent peer-reviewed critiques of the ECI have been published, it is difficult to refute or confirm these concerns (Conte, 2005).

2.1.2 Bar-On’s mixed model of emotional intelligence

Consistent with Goleman, Bar-On’s mixed model approach suggests that EI is comprised of an array of trait and state characteristics, both of which influence an individual’s probability of success. In his model, Bar-On (1997) identifies five areas of functioning related to success: (a) intrapersonal (i.e., emotional self-awareness, assertiveness, self-regard, self-actualization, independence); (b) interpersonal (i.e., interpersonal relationships, social responsibility, empathy); (c) adaptability (i.e., problem solving, reality testing, flexibility); (d) stress-management (i.e., stress tolerance, impulse control); and (e) general mood (i.e., happiness, optimism).
Bar-On uses the Emotional Quotient Inventory (EQ-i: Bar-On, 1997) to measure EI. The EQ-i is a 133-item self-report measure grouped into five higher-order dimensions. Although adequate test-retest reliability \( r = .73 \) (Bar-On, 1997) and acceptable predictive validity \( r = .20 \) (Van Rooy & Viswesvaran, 2004) have been established for the instrument, studies of concurrent validity suggest considerable overlap between the EQ-i and other psychological measures. Review of these convergent and discriminant validity data suggest that many items on the EQ-i pertain to personality attributes (e.g., optimism, emotional stability) (Bracket & Mayer, 2001; Conte, 2005; Dawda & Hart, 2000), so much so that it has been suggested that EI as conceptualized by Bar-On “may be a lower-level primary trait that could be placed below the Big Five in a multistratum model” (Matthews et al., 2004, p. 213).

2.1.3 Mayer and Salovey’s ability model

In contrast to the mixed models explained above, the ability model describes EI as an ability to recognize the meanings of emotions and their relationships, as well as the ability to use emotions to inform cognitive activities (e.g., reasoning, problem-solving) (Mayer et al., 2001). Informed by the aforementioned description, Mayer and Salovey (1997) conceptualized an ability model consisting of the four following skills or branches: (a) Branch 1 (i.e., perception and expression of emotion), which encompasses the ability to identify and express one’s physical states, feelings, and thoughts; (b) Branch 2 (i.e., assimilating emotion in thought), which consists of the ability to use one’s emotions to prioritize thinking in productive ways; (c) Branch 3 (i.e., understanding and analyzing emotion), which encompasses the ability to label emotions and simultaneous feelings, and understand cognitions associated with shifts of emotion; and (d) Branch 4 (i.e., regulation of emotion), which consists of the ability to stay open to regulate emotions reflectively so as to promote emotional and intellectual growth. These branches represent a hierarchical structure whereby it is difficult to manage emotions (Branch 4) if you cannot first understand how your emotions influence your thoughts (Branch 3). Conceptualized in this way, EI is a mental skill or ability that develops over time with training and/or experience.

At the time of writing, two assessment inventories frequently associated with the ability model are the performance-based Mayer-Salovey-Caruso Emotional Intelligence Test, Version 2.0 (MSCEIT V2.0: Mayer et al., 2002) and the perception-based Emotional Intelligence Scale (EIS: Schutte et al., 1998). Early factor analysis calculations for both inventories suggest a factor structure consistent with the ability model of EI, yielding a score for total EI and each of the four branches (i.e., perceiving emotion, assimilating emotion in thought, understanding emotion, managing emotion) (Brackett & Mayer, 2003; Mayer et al., 2003; Schutte et al., 1998). More recently, however, Schutte and colleagues (Riley & Schutte, 2003; Schutte et al., 2002) have suggested that EI is a unidimensional construct with the EIS providing only an overall score of EI.

While research consistently suggests distinction between the MSCEIT and other psychological constructs, discriminant validity of the EIS has been called into question. For example, correlations between the MSCEIT and personality are weak, ranging from \( r = .03 \) to \( r = .28 \) for the Big Five (Brackett & Mayer, 2003; Brackett et al., 2006) and \( r = .04 \) to \( r = .24 \) for the 16PF (O’Connor & Little, 2003). Although the research of Schutte et al.
(1998) suggests nonsignificant correlations between the EIS and four of the Big Five personality constructs, the research of others (Brackett & Mayer, 2003; Brackett et al., 2006) suggests the EIS is not easily distinguishable from either personality or psychological well-being.

2.2 Perceived versus performance measures of emotional intelligence

Alluded to above, the MSCEIT and the EIS represent different types of EI measures, therefore potentially different EI constructs. A performance test like the MSCEIT evaluates individuals’ item responses against objective or predetermined scoring criteria, whereas a self-report or subjective measure such as the EIS asks individuals to judge how good they themselves are at recognizing emotions. In utilizing and/or interpreting measures of EI, five major differences between performance and self-report measures should be considered: (a) performance tests assess actual EI (i.e., maximal attainment), while self-report measures assess perceived EI (i.e., personality traits); (b) performance measures typically require more time to complete, score, and evaluate than self-report measures; (c) self-report measures require respondents to have insight into their own level of EI; (d) self-report measures are susceptible to response and social desirability bias; and (e) self-report measures are more strongly related to personality traits and psychological well-being than performance measures (Matthews et al., 2002; Wilhelm, 2005). It appears then that performance measures of EI are less likely than self-report measures to be influenced by personality and other psychological constructs, thereby providing a more true representation of EI ability.

That said, there is a need to carefully consider not only the interpretation of results emanating from data collected by these diverse approaches, but also the instruments used to collect the data. Specifically, it has been suggested (Austin et al., 2004; Perez et al., 2005; Petrides & Furnham, 2000) that different types of measures assess two different constructs: self-report measures assess trait EI while performance measures assess state EI (i.e., EI ability). These discrepancies between trait and state EI have implications for the utility of the construct in professional practice, and may contribute to the contradictory profiles obtained for the same individuals (Brackett & Mayer, 2003).

The disparities in EI assessment outlined above have implications for future research and application, yet say nothing about the use of electronic data collection. The relative ease and cost effectiveness of online EI data collection may facilitate research, thereby expediting standardization of theoretical and empirical approaches to the study of the construct. While the correlation between PP and online administration of the MSCEIT has been reported at $r = .99$ by creators of the instrument (Mayer et al., 2003), independent confirmation is warranted. Furthermore, examination of the validity of the online version of the self-report instrument (i.e., EIS), and comparison between the self-report and performance measures will inform future EI research. By exploring further the relationship between psychometric equivalence of online and offline EI assessment, the current research contributes to the theoretical and empirical foundation upon which Internet research and EI application can be advanced. The purposes of the current study, therefore, were to: (a) examine the equivalence of online and PP measures of EI, and; (b) provide independent confirmation that the MSCEIT is a more accurate assessment of the ability model of EI, a conclusion that holds true in online format.
3. Methods

3.1 Participants

The sample of participants in the current study consisted of 157 individuals (109 women; 48 men), ranging in age from 19 to 69. For the purposes of this study, online data collected from these individuals were compared to offline data collected from individuals in previous research (Brackett & Mayer, 2003; Schutte et al., 1998). Participants in all three samples were predominantly female and Caucasian. Average age in the current online sample was 31.73 ±11.65, while average age in the two offline samples were 18.93+1.51(F)/19.51±1.17(M) and 29.27±10.23, respectively (Brackett & Mayer, 2003; Schutte et al., 1998).

3.2 Measures

Two EI inventories, both informed by the ability model of Mayer and colleagues, were used to measure the construct. One inventory was used to measure the personality construct.

3.2.1 The MSCEIT

The MSCEIT Version 2.0 (Mayer et al., 2002), a 141-item performance inventory, was used to record scores for total EI as well as each of the four branches (i.e., perceiving emotion, assimilating emotion in thought, understanding emotion, managing emotion). Factor analysis calculations suggest that the MSCEIT has a factor structure consistent with the four-factor model of EI (Brackett & Mayer, 2001). Further analysis suggests a two-week test-retest reliability of $r = .86$. Similarly, results suggest a lack of convergence between the MSCEIT and self-report (i.e., mixed model) EI measures, and discrimination between the MSCEIT and well-being scales as well as Big Five personality measures (Brackett & Mayer, 2001; Conte, 2005). That is, EI as measured by the MSCEIT exists as a mental ability that is distinct from personality variables as well as other mixed measures of EI.

Because the test publisher scores the MSCEIT, we were unable to calculate internal consistencies for our sample (Day & Carroll, 2004). It should be noted, however, that branch score reliability coefficients for consensus scoring have been shown to range from $\alpha = .79$ to $\alpha = .91$ (Mayer et al., 2004).

3.2.2 The EIS

The EIS (Schutte et al., 1998) is a 33-item self-report inventory that assesses the extent to which an individual can identify, understand, harness, and regulate emotions in self and others. Using a unidimensional conceptualization, adequate internal reliability ($r = .87$ to $r = .90$) and test-retest reliability ($r = .78$) have been reported (Schutte et al., 1998). Similarly, meta-analysis results (Van Rooy & Viswesvaran, 2004) indicate that the EIS had higher predictive validity than other EI measures (i.e., EQ-i, ECI). Studies of concurrent validity suggest moderate to strong correlations between the EIS and other personality measures (Brackett & Mayer, 2003; Schutte et al., 1998; Schutte et al., 2002). Examination of discriminant validity data yields contradictory results (Brackett & Mayer, 2003; Ciarrochi et al., 2001; Schutte et al., 1998). For example, the research of Schutte et al. (1998) suggests nonsignificant correlations between the EIS and four of the Big Five constructs while the
research of Bracket & Mayer (2003) suggests that the EIS is not easily distinguishable from either personality or well-being.

Internal consistency reliability for the EIS in the current study was calculated as $\alpha = .89$.

### 3.2.3 Personality

The Mini-Modular Markers (3M40: Saucier, 2002), a 40-item self-report inventory, was used to measure Big Five Personality factors (i.e., extraversion, agreeableness, conscientiousness, emotional stability, openness). For the current sample, internal consistency reliabilities for the five scales of the 3M40 ranged from $\alpha = .74$ to $\alpha = .86$.

### 3.3 Procedure and data analysis

Online data collected in the current study were compared to PP data collected in previous studies.

### 4. Results

#### 4.1 Online versus paper-pencil measures

Calculations of $z$-scores were used to examine differences in correlation coefficients between measures of EI and personality collected via online format in the current study and via PP format in previous research (see Table 1). No statistical differences were detected between electronic and PP data when the MSCEIT was used to assess EI. These results suggest that the relationships between EI (as assessed by the MSCEIT) and personality hold steady when using either electronic or PP data collection.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>.52</td>
<td>.42</td>
<td>.19</td>
<td>-.18</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>3.47***</td>
<td>.86</td>
<td>-.75</td>
<td>-.10</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>.00</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
</tr>
<tr>
<td>Emotional Stability</td>
<td>5.60***</td>
<td>2.98**</td>
<td>.41</td>
<td>.46</td>
</tr>
<tr>
<td>Openness</td>
<td>-1.37</td>
<td>-1.23</td>
<td>-.61</td>
<td>-.57</td>
</tr>
</tbody>
</table>

Note: Current study utilized online data collection while previous studies used paper-pencil collection. 3M40 = Mini Module Markers; EIS = Emotional Intelligence Survey; MSCEIT = Mayer-Salovey-Caruso Emotional Intelligence Test. 

$**p<.01$, $***p<.001$

Table 1. $z$-Score Differences in Correlation Coefficients between Measures of Personality and Emotional Intelligence in Present and Previous Studies.
4.2 MSCEIT vs. EIS

In an effort to compare the psychometric properties of the two disparate online measures of EI, correlation coefficients were calculated (see Table 2). Correlations show a statistically significant yet relatively weak relationship between the MSCEIT and EIS total scores. Correlations also suggest that the EIS has greater convergence with measures of personality than does the MSCEIT. Concomitantly, the MSCEIT displayed divergence from measures of personality, with significant yet weak relationships identified for only the Agreeableness and Openness scales ($r = .27, p < .01; r = .19, p < .05$, respectively). In both cases, as illustrated in Table 1, the EIS had a significantly stronger relationship with those two personality scales. Additionally, when using the EIS three statistically significant differences were identified between electronic data collected in the current study and PP data collected previously. Taken together, these results suggest that the MSCEIT is a more accurate measure of EI ability, a finding which holds true regardless of data collection format.

![Table 2](image)

Table 2. Intercorrelations for Scores on Two Emotional Intelligence Measures and Personality

5. Discussion

The primary purpose of the current study was to examine the equivalence of online and PP measures of EI. No statistical differences were detected between correlations of online and PP administrations of the MSCEIT. Correlations between the MSCEIT and personality also held constant when using online or PP data collection. These results are consistent with previous research conducted by inventory creators who identified uniformity between online and PP administrations of the MSCEIT (Mayer et al., 2003). The finding of psychometric equivalence in the current study is also consistent with studies of other psychological variables (Campos et al., 2011; Lonsdale et al., 2006; Meyerson & Tryon, 2003; Naus et al., 2009; Preckel & Thiemann, 2003), most notably those of Buchanan and Smith (1999) who utilized two different study samples. Taken together, these results support the use of online administration of the MSCEIT to assess EI ability.
A secondary purpose of the current study was to compare the psychometric properties of two disparate online ability-based measures of EI. Tests of convergent and discriminant validity suggest that the EIS conceptualization of EI lacks discriminant validity and is largely indistinguishable from personality constructs (Brackett & Mayer, 2003). It appears, then, that assessment inventories based upon the self-report and trait approaches fail to provide new information about this discrete concept, making it difficult to differentiate between EI and various personality constructs. Similarly, this strong grounding in personality traits contradicts claims that EI is a group of skills that can be learned and developed over time (Goleman, 1995). As such, it may be appropriate to consider the ability model (i.e., MSCEIT) of EI for identifying emotion-based contributions to behavior. Additionally, the self-report nature of the EIS makes it susceptible to social desirability bias (Austin et al., 2004; Schutte et al., 1998), a fact which is exacerbated by a lack of reverse-keyed items (i.e., lie-scale). These findings are consistent with those of other studies using traditional PP administration of the respective assessment inventories (Brackett & Mayer, 2003; Brackett et al., 2006).

5.1 Limitations and suggestions for future research conclusion

Results of the current study support the use of online data collection to assess EI performance. That said, several limitations exist in the current study which should be considered in future research. While tests of convergent and discriminant validity demonstrate the psychometric equivalence of data obtained from electronic and PP administration of the MSCEIT, different samples were utilized. Specifically, online data collected in conjunction with the current study were compared to PP data collected by other researchers on different samples. Although similar methods were used in several studies (Fouladi et al., 2002; Gosling et al., 2004; Meyerson & Tryon, 2003) to validate Internet-based psychological research, these threats to internal validity could be reduced by employing random assignment to online and PP groups or counterbalanced repeated measures designs (Lonsdale et al., 2006).

Similarly, various inventories were utilized to collect the personality data used in this study. Saucier's (2002) 3M40 was used to measure Big Five Personality factors in the current study, while the Revised NEO Personality Inventory was used to measure the Big Five Personality factors in previously conducted studies (Brackett et al., 2006; Brackett & Mayer, 2003). It should be noted that still other measures of personality (i.e., 16 PF, original NEO Personality Inventory) have been used in studies (Caruso et al., 2002; Day & Carroll, 2004; Roberts et al., 2001) examining the psychometric properties of various measures of EI. These disparate methods of assessing personality make it difficult to effectively assess the relationship between EI and personality. As such, researchers should consider coming to an agreement on the most acceptable measure for assessing personality.

6. Conclusion

Research in the area of EI continues to move forward, yet a lack of consistency in the assessment of the construct has slowed progress in the field from both an applied and an empirical standpoint. Given the interest in EI and its implications for performance in a variety of domains (e.g., sport, health, business; Meyer & Fletcher, 2007), the use of an
The Equivalence of Online and Paper-Pencil Measures of Emotional Intelligence

ability model to define and measure EI is deemed most appropriate. In an effort to facilitate evidence-based practice and to advance this line of research (i.e., share and compare results across studies and populations), there is a need to standardize the measurement of EI (Brackett & Mayer, 2003; Meyer & Fletcher, 2007). The discriminant validity of other EI measures (e.g., overlap with personality) and the need for a performance-based evaluation, support the results of the current study which suggest that the MSCEIT V2.0 be used as the criterion for assessment of EI. That said, the financial costs associated with delivery and scoring of PP versions of the MSCEIT V2.0 may have hindered past efforts to standardize assessment of EI. The results of the current study, in conjunction with the relative ease and cost-effectiveness of online data collection, should make it possible for EI researchers to standardize their collection of EI data through use of the MSCEIT V2.0.

7. Acknowledgement

We would like to thank Blake Pindyck for his help in the preparation of this manuscript.

8. References


The Equivalence of Online and Paper-Pencil Measures of Emotional Intelligence


Emotional intelligence is an emerging construct for applied research and possible interventions, both in scholastic, academic and educational contexts, organizational contexts, as well as at an individual level in terms of people’s well-being and life satisfaction. From the presented contributions, it emerges how this volume is characterized by an interest to give an international overview rich of stimuli and perspectives for research and intervention, in relation to a promising variable of current interest, such as emotional intelligence. The goal is that this book further contributes to the affirmation of a particularly promising variable, such as emotional intelligence, which requires a greater interest and attention in both research and application field.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
