



University of Pennsylvania
ScholarlyCommons

Management Papers

Wharton Faculty Research

2015

The Social Perception of Emotional Abilities: Expanding What We Know About Observer Ratings of Emotional Intelligence

Hillary A. Elfenbein

Washington University in St. Louis

Sigal G. Barsade

University of Pennsylvania

Noah C Eisenkraft

University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/mgmt_papers

 Part of the [Management Sciences and Quantitative Methods Commons](#)

Recommended Citation

Elfenbein, H. A., Barsade, S. G., & Eisenkraft, N. (2015). The Social Perception of Emotional Abilities: Expanding What We Know About Observer Ratings of Emotional Intelligence. *Emotion*, 15 (1), 17-34. <http://dx.doi.org/10.1037/a0038436>

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/mgmt_papers/198

For more information, please contact repository@pobox.upenn.edu.

The Social Perception of Emotional Abilities: Expanding What We Know About Observer Ratings of Emotional Intelligence

Abstract

We examine the social perception of emotional intelligence (EI) through the use of observer ratings. Individuals frequently judge others' emotional abilities in real-world settings, yet we know little about the properties of such ratings. This article examines the social perception of EI and expands the evidence to evaluate its reliability and cross-judge agreement, as well as its convergent, divergent, and predictive validity. Three studies use real-world colleagues as observers and data from 2,521 participants. Results indicate significant consensus across observers about targets' EI, moderate but significant self-observer agreement, and modest but relatively consistent discriminant validity across the components of EI. Observer ratings significantly predicted interdependent task performance, even after controlling for numerous factors. Notably, predictive validity was greater for observer-rated than for self-rated or ability-tested EI. We discuss the minimal associations of observer ratings with ability-tested EI, study limitations, future directions, and practical implications.

Keywords

emotional intelligence, social perception, observer ratings, self-ratings, ability tests, social relations model

Disciplines

Management Sciences and Quantitative Methods

Running Head: SOCIAL PERCEPTION OF EMOTIONAL INTELLIGENCE

The social perception of emotional abilities:

Expanding what we know about observer ratings of emotional intelligence

Hillary Anger Elfenbein

Washington University in St. Louis

Sigal G. Barsade

University of Pennsylvania

Noah Eisenkraft

University of North Carolina – Chapel Hill

KEY WORDS: Emotional Intelligence, social perception, observer ratings, self-ratings, ability tests, Social Relations Model

Abstract

We examine the social perception of emotional intelligence (EI) through the use of observer ratings. Individuals frequently judge others' emotional abilities in real-world settings, yet we know little about the properties of such ratings. This paper examines the social perception of EI and expands the evidence to evaluate its reliability and cross-judge agreement, as well as convergent, divergent, and predictive validity. Three studies use real-world colleagues as observers and data from 2,521 participants. Results indicate significant consensus across observers about targets' EI, moderate but significant self-observer agreement, modest but relatively consistent discriminant validity across the components of EI, and significant predictive validity of observer ratings in work and task performance domains, even after controlling for cognitive intelligence, personality, trait affect, observer liking, and demographic characteristics. We discuss the poor associations of observer ratings with ability-tested EI, study limitations, future directions, and practical implications.

The social perception of emotional abilities:

Expanding what we know about observer ratings of emotional intelligence

The public interest and academic research in *emotional intelligence* (EI) is at least partly inspired by the commonsense notion that those around us vary in their emotional skills. It is relatively easy to point out examples and cautionary tales of EI in literary works as well as real life (e.g., Oatley, 2004). Indeed, among the reasons for the prominence of the EI construct, Averill (2004) argued, “is that it has an easily recognizable grain of truth. We all know people who are emotionally adept” (p. 228). In the present paper, we focus on this phenomenon itself—that is, people’s beliefs about other people’s emotional abilities. We take a social perception lens to the construct of emotional intelligence and, in doing so, attempt systematically to map a greater understanding of observer ratings.

We are interested in the social perception of emotional abilities as an epiphenomenal construct—that is, judgments of EI are theoretically important in their own right because they exist in our minds, not necessarily because they are accurate or valid for use as a measurement device. People continually make these judgments, and these judgments subsequently have meaningful interpersonal consequences. As Funder and West (1993) argued, interpersonal perceptions “are the basis of such common, consequential decisions as whom to befriend, trust, avoid, hire, promote, release from prison, or elect as president” (pp. 457-458). Because judgments of other people’s emotional abilities are involved in these types of important decisions, it is worthwhile to understand them better. Indeed, Roberts and colleagues (2001) argued that in the field of emotional intelligence “validation studies of this type [i.e., observer perceptions]...[are] in urgent need of detailed investigation” (p. 201). The present investigation attempts to fill this gap.

Social perception

In presenting this work, we first observe that a large body of research about how observers judge other characteristics in the social environment. Researchers have also long studied social perception in

the realm of behavioral tendencies—that is, personality. Given the theoretical foundation that defines personality as consistency over time in an individual's behaviors when that individual is placed in the same situation (e.g., Mischel & Shoda, 1995), the judgment of internal personality traits is based on observing people's outward behavior. Indeed, leading personality trait theorists for decades have treated observer reports as a gold standard that can be used for validating self-reports, and vice versa (Kolar, Funder, & Colvin, 1996). Funder (1991) went as far as to argue that “the single best method of trait assessment is observer report....For social traits at least, it is hard to imagine a higher court of evidential appeal that could over-rule observers' judgments, assuming the observers have had ample opportunity to observe the target's behavior in daily life” (p. 35). Empirical evidence has shown strong cross-observer and self-observer convergence for core personality traits (Funder & Colvin, 1997) as well as for trait affect (Watson & Clark, 1991). Correlations are typically in the .20-.40 range for the big five personality traits, with highly acquainted raters towards the top end of the range (John & Robins, 1993). That said, there is evidence that the properties of social perception are less clear for traits that are socially desirable. In particular, there is no apparent agreement between self and observer reports of wisdom (Redzanowski & Glück, 2013), which might result from challenges with either or both sources.

There is also a stream of research documenting the perception of abilities other than EI. The association between IQ tests and observer ratings tends to be significant, even if modest in magnitude (Mabe & West, 1982; Paulhus & Morgan, 1997). Although abilities are internal to the person, they tend to be displayed in outward behaviors that are themselves observable. Informed observers can generally perceive cognitive intelligence because it is connected with cognitively intelligent choices and outcomes—one can view chess games, recognize who tends to be logical in conversations, and observe an individual's acceptance to particular universities and job positions. That is, there are traces of information in the environment regarding abilities, even if this information can be imperfect. Likewise, close observers might attempt to sample EI from multiple observations and to view the behavioral

consequences of this ability in order to judge an individual's effectiveness in emotional domains—whether or not these observations are valid

The social perception of EI could also be consistent with ability models of emotional intelligence. Theorists now recognize the ability model as the most authoritative approach to EI (Mayer et al., 2008; Matthews, Roberts, & Zeidner, 2004; Rivers, Brackett, Salovey, & Mayer, 2007). This model begins with the larger concept of *intelligence*, the various definitions of which typically involve a person's ability to "deal effectively his [or her] environment" (Wechsler, 1944, p. 3, see also Roberts, Zeidner, & Matthews, 2001). Top theorists of emotional intelligence argue that the criteria for what it means to deal effectively on an emotional level are socially shared. With this reasoning, Mayer, Salovey, Caruso, and Sitarenios (2003) argued that using consensus scoring as a method of determining correct answers for their test is appropriate because their "model of EI hypothesizes that emotional knowledge is embedded within a general, evolved, social context of communication" (p. 98) that allows even untrained samples to define correct responses. Empirically, Mayer et al. (2003) find that there is a high level of agreement regarding the correct answers to the items on their ability test, as well as a high rate of agreement between the responses of researchers and those of the general public used for the consensus coding. This argument could be extended to suggest that, if untrained observers can perceive skilled responses on a test of emotional intelligence ability, then they might also be able to perceive to what degree other people typically make these skilled responses in actual situations.

Taken together, we argue that it is worthwhile to conduct studies of observer ratings of emotional intelligence, and add to the body of knowledge about social perception in general and about EI in particular. Our goal in the current paper is to expand the empirical basis of evidence to understand the properties of observers' judgments about other people's emotional abilities.

Evaluating the observer report of EI

The goal of this paper is to provide a comprehensive empirical investigation of the social

perception of EI, through a large-scale investigation of observer ratings. This investigation includes the key psychometric properties involved with evaluating validity, which are detailed below.

Consensus among judges. In attempting to unpack the meaning of observer perceptions, it is important to know whether observers agree. Note that this condition is necessary for validity but not sufficient—that is, in theory observers could completely agree yet be completely wrong, as in the case of using false stereotypes or inaccurate cues for their judgments (Hammond & Stewart, 2001). However, a lack of agreement would preclude validity because no consistent signal emerges.

Some evidence already exists for moderate consensus in judging emotional abilities. To date, a handful of EI studies have included observer ratings—often, but not always, in an attempt to validate the self-report instruments that were of primary interest to the investigations (e.g., Law et al., 2004). For this reason, most of these studies used a single judge or used multiple judges without reporting cross-judge agreement (e.g., Law et al., 2004; van der Zee, Thijs, & Schakel, 2002). Kellett, Humphrey, and Sleeth (2006) reported intra-class correlations for cross-judge agreement of .32 for emotional expression ability and .23 for emotion perception ability in their study. Likewise, researchers have reported convergence between the EI ratings made by participants' peers and supervisors ($r=.34$, Law et al., 2004; $r_s=.23-.30$; Lopes, Grewal, Kadis, Gall & Salovey, 2006). Charbonneau and Nicol (2002) reported that inter-rater agreement (r_{wg}) on interpersonal EI was greater than .70 for most of their sample. Further, they reported that peer and supervisor ratings correlated at $r=.22$, which is moderate but suggests some underlying signal that observers detected in common. These values compare moderately with the relatively good convergence between peer and supervisor ratings of $r=.48$ found in Harris and Schaubroeck's (1988) meta-analysis of workplace job performance. The current paper expands this evidence by examining every factor separately from two different four-factor models of EI, using advanced statistical techniques with the authoritative Social Relations Model for social perception (Kenny, 2004), and by increasing greatly the existing body of data with over 2500 participants.

Convergent validity. Although observer reports of emotional abilities are important epiphenomenal constructs in their own right, it is worthwhile to examine their convergent validity with respect to well-established ability tests of EI. Given that, as discussed above, ability models of EI argue that even untrained observers can recognize what behaviors are considered emotionally intelligent (Mayer et al., 2003), it is also possible that these same observers can recognize the extent to which close others typically produce such behavior in daily life. There is some limited existing evidence for significant associations between observer ratings and ability tests of emotional skills in nonverbal communication ($r=.20$; Rosenthal et al., 1979) and interpersonal sensitivity ($r=.48$; Costanzo & Archer, 1989). Lopes et al. (2006) found convergence between an ability test of overall EI with peer and supervisor reports of trait EI. However, other studies have shown no association among ability tests and observer reports, such as Carney and Harrigan's (2003) investigation of both emotion perception accuracy and social situation understanding.

Divergent validity. A longstanding concern among theorists of emotional intelligence is the criticism that EI might merely be a repackaging of old wine in new bottles (Matthews, Zeidner, & Roberts, 2002; Mayer, Caruso, & Salovey, 1999). As such, it is important to examine any concept related to EI with other constructs that have positive valence. We begin with liking, for its potential halo effect that could cast a positive glow across ratings of other desirable characteristics, and we include more general positive regard for the same reason. We also include observer ratings of emotional stability, i.e., low levels of neuroticism, for the potentially confounding nature with this other emotion-relevant personality trait. In the little research conducted so far on this topic, Song et al. (2010) found only a small association, $r=.12$, between liking and observer-rated EI. We add to this evidence with an expanded list of potential confounds and confirmatory factor models using large numbers of participants.

Predictive validity. The predictive validity of EI has been considered a 'holy grail' in the field, and quests to find it have been met with tempered success (Matthews et al., 2002; O'Boyle, Humphrey,

Pollack, Hawver, & Story, 2011). As such, it is worthwhile to examine the potential predictive validity of observer-rated EI. There is some initial evidence that observer-reported EI does predict important outcomes. Law et al. (2004) measured observer reports as a supporting measure in validating their self-report scale of EI, and they presented evidence that observer-rated EI had substantial predictive validity for job performance—even controlling for the Big Five personality dimensions. Indeed, the predictive validity of observer-rated EI was higher than that of self-reported EI in their study. Law et al. (2004) also showed that parent-rated EI had predictive validity for life satisfaction. Song, Huang, Peng, Law, Wong, and Chen (2010) found that EI judgments from college roommates predicted academic performance and the quality of social interactions above and beyond cognitive ability and personality. EI judgments from college students' friends predicted both grades and their participation in social activities (van der Zee et al., 2002). Sala (2005) reviewed evidence showing that observer ratings on the emotional competence inventory predict a range of criteria such as job performance across several industries, feelings of life success, perceptions of leadership, and lower levels of stress (see also Boyatzis & Sala, 2004). Our study extends the body of evidence in the workplace domain by examining the predictive power of observer-rated EI beyond that of a wide range of control variables, including conventional intelligence and existing personality traits.

Self-observer agreement. It is also worthwhile to explore the extent to which observer reports converge with self-perceptions. In doing so, it is important to emphasize that self-reported EI cannot be taken at face value as a veridical construct, given the challenges to developing accurate self-awareness about one's emotional abilities. Abilities can be self-reported the most accurately in those domains for which there are 'gold standards' for testing and for which people receive explicit feedback about where they stand. This includes mathematical and verbal intelligence as well as athletic abilities (Kuncel, Credé, & Thomas, 2005; Mabe & West, 1982). Because this criterion is not the case for EI—where we live in a 'white lie' society (Swann, Stein-Seroussi, & McNulty, 1992) that provides relatively little clear feedback

about our interpersonal skills—the accuracy of self-reported EI is likely to be imperfect. Expecting individuals to report about their own emotional skills may be asking them to tell more than they could know (Nisbett & Wilson, 1977). Further, somewhat paradoxically, those without EI may lack the metacognitive skills to be aware of their low EI (Brackett, Rivers, Shiffman, Lerner, & Salovey, 2006), adding to arguments against taking the self-report of EI at face value. Some theorists have even referred to trait models of EI as “emotional self-efficacy” (Petrides, et al., 2007, p. 151; Tett, Fox, & Wang, 2005, p. 859), which suggests that self-reported emotional traits represent one’s priorities, motivations, and hopes for emotional processes as much as skill per se. Thus, like observer reports, self-reports of EI are also epiphenomenal—theoretically important in their own right because they exist in our minds. In order to examine the same social perception both from the lens of the observer and the target, self-observer agreement is of interest in the domain of EI.

There are two distinct theoretical mechanisms that could produce significant self-observer agreement. First, there can be some accurate signal that makes its way through the noise of both self-awareness and observer perception. People can observe their own behaviors to see whether we achieve their emotionally desired ends, and other people can give their feedback along the way. Second, through self-verification processes, people can help to shape others’ perceptions in order to persuade them, however subtly, to view them the way they view ourselves (Swann, Chang-Schneider, & McClarty, 2007).

There is already some evidence in support of significant—even if moderate—associations between observer reports and self-reported emotional intelligence. Law et al. (2004) demonstrated good self-observer agreement for total EI rated by parents, workplace observers, and supervisors, with coefficients from $r=.28-.41$. These values are on par with the agreement that the same authors found for the big five personality traits. Other reported self-observer correlations in emotional domains have been $r=.20$ for emotional self-awareness, $r=.43$ for emotional self-control, and $r=.29$ for empathy Sala (2005); $r=.33$ for emotional self-awareness (Bar-On & Handley, 2003), and $r=.13$ for emotion perception accuracy

(Carney & Harrigan, 2003). In research using measures from the trait EI perspective, there have been reports of $r=.09$ for interpersonal EI (Charbonneau & Nicol, 2002), $r=.18$ for an overall EI inventory (Jordan & Ashkanasy, 2005), and $r=.12$ for empathy rated by unacquainted observers (D'Augelli, 1973). One report reported self-observer agreement correlations as high as .56-.65 (van der Zee et al., 2002), whereas others tend to be more moderate, e.g., $r=.26$ to .35 in Libbrecht, Lievens, and Schollaert (2010). This general picture of self-observer agreement across emotional skills is in line with the level of self-peer agreement of $r=.24$ for job performance based on Harris and Schaubroeck's (1988) meta-analysis. When considering observer ratings of traditional cognitive intelligence, self-other agreement is typically in the range of low .30s for community samples and rarely greater than .30 for college students (Paulhus, Lysy, & Yik, 1998). As with observer agreement, we expand existing evidence with a comprehensive examination of two different four-factor models.

Measurement equivalence between peer and self-ratings. It is worthwhile to expand the evidence for evaluating the measurement equivalence between self and observer reports of EI. Existing data have shown that these two types of measurement are largely equivalent in measuring the same underlying construct, albeit with some leniency bias for self-ratings (Joseph & Newman, 2010; Libbrecht et al., 2010).

The current studies

Below we report the results of three studies conducted with real-world colleagues to examine the properties of social judgments of emotional intelligence. Study 1 offers a first examination of observer reports within a large sample of teams of full-time MBA students who worked closely together, using observer reports adapted from a well-validated self-report questionnaire. Study 2 extends these findings using observer reports of EI and performance reviews from individuals' workplace colleagues and supervisors, respectively, and expands the evidence by adapting observer reports from a different EI instrument. Study 3 attempts to address potential limitations of the first two studies by demonstrating that observer ratings of EI are distinct from liking, positive regard, and emotional stability, includes an

ability test of EI to evaluate concurrent validity, and tests the predictive validity of observer rated EI while controlling for liking and positive regard. Taken as a whole, these studies present the most comprehensive test to date for evaluating the social perception of emotional intelligence.

Study 1

We take a social perception approach to the study of EI and, for this reason, use the authoritative conceptual model developed for the study of interpersonal perception, Kenny's (1994) Social Relations Model (SRM; see also Kenny & LaVoie, 1984).

The Social Relations Model describes how all interpersonal ratings—such as those of EI—are a function of four mutually exclusive and exhaustive components. First, *target effects* consist of consensus across raters, in that the group as a whole systematically rates some people as having higher versus lower EI. This is our most important coefficient, and it is the operational definition of observer-rated EI. Second, *rater effects* consist of response bias on the part of raters who vary in the average scores that they assign when judging others' EI. This is a matter of leniency bias. Third, *relationship effects* consist of dyadic perceptions whereby some people uniquely find particular others to be very high or low in EI—even after controlling for the perceiver's general rating of others and the target's typical rating by others. The fourth component of the SRM is measurement error, which results from imperfect inter-item correlations within the questionnaire. For the purpose of this study, we use the SRM first to examine the extent to which interpersonal ratings of EI show consensus across raters. We use it second to construct variables that represent each individual's aggregate ratings from multiple teammates.

To distinguish these components from each other statistically, researchers use designs in which each participant rates multiple targets. We use the most common of these designs, the *round robin*, in which each member of a group rates each other member of the group (Kenny & LaVoie, 1984). The Social Relations Model offers several advantages to the study of observer-rated EI. First, it accounts for the inherently interdependent structure of ratings data. Notably, in the case of the round robin, there are

systematic missing data in that participants do not provide 'observer' ratings of themselves. This is noteworthy because the conventional analysis of such data would be vulnerable to artifacts. For example, an individual who gives everyone high ratings would appear low in EI because she missed the chance to have such a generous grader, and likewise an individual who gives everyone low ratings would inadvertently inflate his own score. This makes it inappropriate to use statistics relying on the independence of data points, notably the within-group interrater index r_{wg} —a measure of agreement in the case of interchangeable raters of a single target (James, Demaree, & Wolf, 1993). However, the SRM corrects for these factors and for other biases that can influence observer- and self-reports. The SRM models bias explicitly in its algorithms and thereby can correct for it when reporting self-observer agreement. Thus the positive features of the Social Relations Model make it the optimal approach for expanding evidence to evaluate observer ratings of EI.

Method

Participants

This study sampled full-time students in a large Master of Business Administration (MBA) program on the East Coast of the United States. Students completed the study measures as part of their required Organizational Behavior course, and the Institutional Review Board (IRB) approved these data to be used additionally for research purposes after removing individual identifiers. In their first-year coursework, $N=998$ were assigned into student teams of five to six members who completed assignments together across all courses (response rate 93%, for $N=931$ individuals and 186 teams; mean $n=5.31$ individuals per team, $SD=0.72$). Before participants arrived at the university, team composition was determined by a process that ensured diverse representation in terms of sex, national origin, and the nature of previous work experience. This process reduced, but did not completely eliminate, the likelihood of prior acquaintance. We collected observer ratings from close colleagues in the context of these teams.

We chose this population for a number of reasons. First, the observers were well-acquainted after working closely together full-time for approximately two months, across a range of settings—quantitative and qualitative course assignments, including problem sets, business case analyses, and a major group project requiring extensive coordination—which had real consequences. Observer ratings tend to have greater agreement with increasing acquaintance, particularly with opportunities to view the attributes in question (Funder & Colvin, 1988; Watson & Clark, 1991). Second, ability measurement theorists argue that empirical data should be collected among individuals who are relatively comparable along attributes such as age, sex, education level, and other factors that might influence performance (Carroll, 1993). Third, interpersonal judgments tend to have the greatest consensus among raters who have been exposed to each other in group settings—giving all observers experiences in common vs. separate one-on-one contact (Kenny, 1994). Fourth, for observer ratings of personality, Costa and McCrae (1992) found that four appeared to be the optimal number of raters, with diminishing returns for larger numbers. Last, the individuals collaborated on tasks that had real consequences, which allowed them to form opinions about ecologically valid outcome variables, notably work performance.

Measures

Participants completed measures via the Internet in two waves. First, no later than the second day of the course, participants completed all the individual difference measures except as specified below. Second, at the end of the course—approximately eight weeks later—participants completed observer ratings of EI and performance, as well as the remaining individual difference measures. In the case of missing data on EI or performance, analyses below used listwise deletion. In the case of missing data on control measures, analyses used mean substitution.

Self-ratings of EI. Among the self-rating scales that have been developed to measure EI, some of the most extensive published validity data has been documented for the Wong and Law Emotional Intelligence Scale (WLEIS; Wong & Law, 2002). The 16-item WLEIS is based on Davies, Stankov, and

Roberts' (1998) model of EI derived from factor analysis of most EI-related scales available at the time. This model includes the following factors: (a) *Self-Emotions Appraisal* (SEA), which is intrapersonal awareness of one's own emotional states; (b) *Others'-Emotions Appraisal* (OEA), which is interpersonal awareness of other people's emotional states; (c) *Use of Emotion* (UOE), which is using one's own emotional states to bring out one's best motivation and achievement; and (d) *Regulation of Emotion* (ROE), which is the ability to control one's own emotions. The authors selected four questionnaire items for each factor on the basis of factor loadings and correlations with criterion variables across multiple field samples. Law et al. (2004) presented further validity data on the scale in the form of confirmatory factor analysis of their model, convergent validity with respect to reports from peers, parents, and supervisors, and incremental predictive validity of criteria such as life satisfaction and job performance while controlling for self-reported personality traits. Even with this validity evidence, it is worth noting that the UOE scale has items that are double-barreled and that have construct overlap with self-efficacy, and results below are interpreted in this light.

Self-ratings of EI were completed in the same session as observer ratings, which took place at the end of the course to maximize acquaintance. Instructions for these evaluations emphasized their confidential nature—i.e., that they would be shared with targets only in aggregate—which is particularly important given findings that self-ratings of ability are more accurate with such assurances (Mabe & West, 1982).

Observer ratings of EI. In keeping with the convention of converting self-reported personality scales into observer-reported versions (e.g., McCrae & Costa, 1987), we adapted the Law et al. (2004) scale by replacing the word "I" with the teammate's first name, the word "me" with "him/her", and the word "my" with "his/her". For questions that were highly intrapersonal, we also added "seems to" before the verb in order to legitimize taking one's best guess in rating another person's internal functioning.

Because SRM analyses cannot accommodate missing data (Kenny, 1998), we excluded from analyses any

individual without complete observer ratings, which affected 6.5% of participants.

Performance ratings from teammates. For the purpose of developmental feedback, team members rated each other's performance at the end of the semester. Feedback was provided after the course ended in an aggregated form to preserve the privacy of individual respondents. They rated each colleague's performance as a *Teammate* ("How effective a member of your team was [name]?"), *Individual* ("How effective an individual contributor was [name]?"), and *Leader* ("How effective a leader was [name]?"), using scales from 1 (not at all effective) to 5 (extremely effective).

Demographic background. Given that EI and/or performance ratings could be differentially related to sex (e.g. Brackett, Mayer & Warner, 2004; Brackett et al., 2006), age (Mayer, Caruso, & Salovey, 1999), and sharing the culture's dominant first language (Downs, Strand, & Cerna, 2007), we included these factors as control variables in predictive validity analyses.

Personality scales. To establish the divergent validity of EI measures with respect to conventional personality traits, participants completed several self-report scales. We assessed the traits of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness using John and Srivastava's (1999) Big Five Inventory (BFI). Tellegen's (1982) multidimensional personality questionnaire (MPQ) measured Trait Positive Affectivity (Trait PA) and Trait Negative Affectivity (Trait NA). Trait PA and NA were particularly worthwhile to include in light of the association between workplace outcomes and trait affectivity (e.g., Côté, 1999; Staw & Barsade, 1993; Staw et al., 1994). Due to time constraints, the MPQ appeared in the second wave of measures. We note that the expansive use of personality characteristics is worthwhile in research that includes self-reported EI, in that it tends to be confounded with a wide range of other factors (Joseph, Jin, Newman, & O'Boyle, 2014).

Cognitive intelligence. As a measure of cognitive intelligence, we used participants' verbal, quantitative, and total scores for the Graduate Management Admissions Test (GMAT). Cognitive intelligence—particularly verbal intelligence—is a worthwhile control variable when establishing the

predictive validity of emotional intelligence (Mayer et al., 2008). Although values were self-reported, meta-analytic evidence shows very high associations between students' self-reports and actual standardized test scores (Newman & Lyon, 2009).

Results

Supplementary Table S1 summarizes descriptive statistics, reliability levels, and bivariate correlations for all study measures.

Consensus among judges

Table 1 summarizes the output of a series of Social Relations Models. As discussed above, the SRM provides estimates for the amount of variance in colleagues' ratings that can be attributed to target, the perceiver, the idiosyncratic relationship between the target and perceiver, and measurement error. In the current model, target variance indicates the degree of consensus among judges in the EI of each focal individual. The magnitude of variance can be interpreted akin to an R^2 value in conventional statistical models. In support of agreement among judges, the level of target variance for each of the four components of EI ranged from 9% to 25%, and this variance was statistically significant and exceeded the 5% threshold suggested for interpreting SRM effects as meaningful (Kenny, 1994).

Although this exploration is focused on target variance, which indicates cross-judge consensus, for completeness with respect to the SRM and for the sake of exploring other properties of observer-rated EI, we also report effect sizes for rater variance. In the present design, in which individuals recorded impressions of each other's EI using a rating scale, rater variance can be interpreted as a type of artifact or response bias. It refers to the extent to which individual raters use higher versus lower mean values in assigning all of their observer ratings. It is worth noting from Table 1 that this response bias is the largest total component, and it is larger than the degree of consensus—i.e., target variance—in most cases. This response bias is controlled in analyses below via the SRM.

Self-observer agreement

There was significant agreement between observer and self-ratings, as shown in Table 2 presents correlations among these variables. Self-observer agreement is positive and significant for all four branches of EI: SEA ($r=.08$, 95% CI .02-.15), OEA ($r=.14$, 95% CI .07-.20), UOE ($r=.24$, 95% CI .18-.31), ROE ($r=.25$, 95% CI .19-.31), and overall EI ($r=.15$, 95% CI .08-.21).

Convergent and divergent validity

In mapping out the properties of observer-rated EI, the data presented in Table 2 provide the information necessary to use Campbell and Fiske's (1959) classic multi-trait multi-method (MTMM) approach to explore convergent and divergent validity. Using MTMM, researchers assess several distinct constructs using two or more methods, and they examine several criteria. First, scores for a particular construct should be similar across different measurement strategies—that is, the self-observer agreement hypothesized above. Second, scores for the same construct should be more similar than scores for distinct constructs. In their large-scale study of trait emotional intelligence, Law et al. (2004) conducted an MTMM analysis that contrasted overall EI with conventional personality traits. They found moderate evidence for discriminant validity, yet also the presence of common method bias in that self-reports tended to converge along a dimension of social desirability. In the present case, the variables are four factors contained within EI, and the analysis explores whether there is MTMM evidence among these four factors—instead of contrasting EI with personality traits. That is, the analysis examines whether the individual components of EI show convergence when measured using different methods and whether they show divergence with respect to each other. Note that Study 1 used only two methods to assess EI, and therefore did not have sufficient degrees of freedom to make a formal test of these criteria using structural equation modeling vs. visual inspection (J. Edwards, private correspondence).

Along these lines, it is worth noting that this MTMM is a highly conservative and stringent test, in that the analysis is typically applied to variables that are theoretically more distinct than four factors within a single construct, as in the present case. The veridical relationships among the four factors of EI

would make it more difficult to establish construct validity. The possibility of halo effects or stereotypes about who might have higher vs. lower EI should also influence all four factors equally, and should make it harder to establish construct validity. Theories posit that the multiple factors within higher-order intelligences should have a positive manifold—that is, non-negative correlations with each other (Gutman & Levy, 1991; Roberts et al., 2001)—and so we still expect positive correlations among the factors of EI, which represent theoretical substance rather than mere methodological artifact. Thus, the MTMM analyses presented below are a highly conservative test of the construct validity of observer-rated EI.

These values provide evidence, albeit modest, for the criteria outlined by Campbell and Fiske (1959). First, the reliability coefficients— listed in the first column of Table 2—are the largest values in any associated row or column. Second, as reported above, the coefficients for self-observer agreement, although not large, are significantly different from zero. Third, these data indicate some discriminant validity. The criterion to apply is that the convergent correlation (i.e., self-observer agreement) should be higher than any of the other values in its row or column within the 4x4 heteromethod block—i.e., the values in bold font should be higher than the values appearing in underline. This criterion is met with the exception of Self-Emotions Appraisal. Fourth, there appear to be common method variance challenges, in that the italicized numbers are generally larger than the degree of observer-self agreement. This suggests that raters are imperfect in their ability to distinguish among the four factors within the umbrella of EI. However, also as discussed above, this could indicate a veridical property of the EI construct, namely that there should be non-negative correlations among multiple factors of intelligence (Gutman & Levy, 1991). This positive manifold should apply both to observer and self- ratings. Given the findings described above, this MTMM analysis of emotional intelligence ratings shows fair construct validity.

Predictive validity of observer ratings

Research in emotional intelligence has been called upon to demonstrate the power of EI measures to predict effectiveness for real-world outcomes of interest. Table 3 summarizes the results of partial correlations that examine observer ratings of effectiveness as predicted by observer-rated and self-rated EI. For analyses of peer-rated EI, we used bootstrap simulations (Efron, 1979; Efron & Tibshirani, 1993) to eliminate same-source bias in that both EI and performance were rated by the same observers (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). In each of 10,000 bootstrap iterations, we randomly sampled a focal individual, one teammate to provide observer ratings of performance, and a separate teammate to provide observer ratings of EI. All of the reported partial correlations control for demographic background, cognitive intelligence, personality, and trait affectivity, in order to examine the incremental predictive validity of EI above and beyond those factors. As seen in Table 3, higher observer-rated EI significantly predicts higher effectiveness ratings, which held across ratings of effectiveness as a teammate, individual, and leader. It is worth noting that there were significant zero-order correlations between self-rated EI and performance, as seen in Supplementary Table S1 but—unlike observer-rated EI—for self-rated EI these associations became non-significant after control variables were included, as seen Table 3. Baseline models including control variables appear in Supplementary Table S2.

Measurement equivalence of self and observer ratings

Table 4 presents the results of analyses of measurement equivalence between self and observer ratings of EI. Establishing measurement equivalence involves comparing a series of nested multigroup confirmatory factor analysis models, which provide increasing constraints on the comparison of CFA solutions for self-reported data vs. observer-reported data. The baseline model (configural invariance) examines to what extent there is the same pattern of factor loading across both observer ratings and self-ratings, while allowing their values of factor loadings to differ. The nested weak invariance model adds a constraint of equal factor loadings. The strong invariance model adds a further constraint of equal intercepts. The final model adds the constraint of equal mean values for the two types of measurement.

The accepted criterion for comparing the relative fit is the change in fit using the CFI index, for which a value smaller or equal to -0.01 indicates there is no significant difference between the two measurements (Cheung & Rensvold, 2002; Joseph & Newman, 2010). This criterion was met for each nested model, indicating support for measurement equivalence.

Discussion

This study evaluated systematically the social perception of emotional intelligence via observer ratings. It is the first to examine the topic using the Social Relations Model (Kenny, 1994), which is based in the psychology of interpersonal perception, and is the authoritative method for examining this topic.

The results provided some evidence for positive psychometric properties, albeit of modest effect sizes. There was consensus across raters in their judgments, which suggests that there is likely to be some kind of underlying signal about EI upon which observers agreed—whether or not that signal is valid. Second, there were significant levels of self-observer agreement. Third, the criteria outlined by Campbell and Fiske (1959) for multi-trait multi-method (MTMM) models of construct validity were modestly met. In a test of predictive validity, higher observer-rated EI predicted better performance ratings, even when controlling for cognitive intelligence, personality, and trait affectivity. The bootstrap methodology corrected for the potential artifacts of common method bias due to response tendencies or idiosyncratic liking between individuals. It is noteworthy that the effect sizes appeared to vary across the four EI components represented in the study. Notably, the Self-emotions appraisal branch had lower consensus among raters and lower self-observer agreement than the other branches. We speculate that this component—referring to the intrapersonal awareness of one's own emotional states—may be associated with less visible cue that observers may use to form their judgments. Last, self and observer ratings appear to satisfy the criteria for measurement equivalence.

Study 2 expands the body of evidence in several ways. First, it increases the generality of the findings beyond the specific questionnaire items used, by collecting observer reports using questions adapted from a different emotional intelligence self-report instrument. Study 1 used the widely cited Wong and Law (2002) scale. However, this scale does not include all of the components of EI that have been covered in theoretical models—notably clarity in the nonverbal expression of emotion and managing the emotions of others (e.g., Mayer, Salovey, & Caruso, 2000; Tett et al., 2005). Expanding generality is also worthwhile, as described above, because the use of emotion scale has questions that are double-barreled or overlapping with the self-efficacy construct. Second, student teams may not generalize directly to other types of observers. The teams in Study 1 interacted intensively and had tangible course grades at stake for students' teamwork, and the individuals were MBA students who had significant job experience and concern about maintaining a positive professional reputation while engaging in tasks where performance relatively visible to their peers. Even so, Study 2 increases external validity by sampling ratings from members of participants' real-world full-time workplaces, and by soliciting criterion variables from their actual supervisors. Further, we examine the influence of observer-rated EI on not only traditional job performance, but also on other workplace factors that have been theorized as relevant to ability EI, such as transformational leadership and organizational citizenship behaviors (Côté & Miners, 2006; c.f., Day & Carroll, 2004). Past work on self-reported EI has found positive associations with ratings of leadership and organizational citizenship, albeit substantially more so for self-reported vs. ability measures of EI (Carmeli & Josman, 2006; Harms & Credé, 2010; Modassir & Singh, 2008). As such, it is worthwhile to examine the associations between these factors and observer-reported EI.

Method

Participants

Students enrolled in a part-time MBA program in the Midwest United States completed the study measures as part of a leadership course, and consented for these data to be used additionally for research purposes. Students in this program typically maintained full-time employment in professional-level positions at corporations that helped to sponsor their part-time studies, and most had leadership roles with at least one supervisee. Participants solicited peer ratings and supervisor ratings ratings from colleagues at their workplace. A total of N=135 focal individuals completed all measures. Including 328 supervisors (average of 2.4 each) and 828 peers (average of 6.1 each), a total of 1,291 people contributed data to the study.

Participants were asked to invite only knowledgeable others to rate them. This criterion is important to increase the validity of ratings (John & Robins, 1993; Kolar, et al., 1996), and so we included data only from observers who responded to the question “How closely have you worked with this person?” with 3 or higher on a 5-point scale (1 = a little, 5 = very closely). This criterion eliminated 74 raters, for a total of 961 raters whose data are included below.

Measures

Before the course began, participants completed a 1.5-hr battery of self-rated and ability measures, and they sent invitations to peers and supervisors to complete 5-10 minute questionnaires each.

Self-ratings of Emotional Intelligence. We used Brackett et al.’s (2006) 19-item Self-Report Emotional Intelligence Scale (SREIS), which was written and validated to correspond to the Mayer and Salovey (1997) Emotional Intelligence model. This model includes four branches: (a) *Perceiving Emotion*, which is the ability to identify emotions in oneself and others; (b) *Using Emotion*, which is the ability to harness feelings towards cognition; (c) *Understanding Emotion*, which is the ability to use language and thinking to analyze emotion; and (d) *Management of Emotion*, which is the ability to regulate one’s own emotions and to influence the emotional states of other people. Brackett et al. (2006) validated the SREIS’s construct

validity, factor structure, and found moderate but significant convergence with a frequently used emotional intelligence ability test, the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT, Mayer, Salovey & Caruso, 2002). Note that the Use of Emotion branch contains two double-barreled questions that are worded with socially-desirable responses in the direction that indicates low EI (“I am a rational person and I rarely, if ever, consult my feelings to make a decision,” and “I am a rational person and don’t like to rely on my feelings to make decisions.”). For this reason, we provide analyses below that include and exclude these two items.

Observer ratings of emotional intelligence. As in Study 1, we adapted the self-report scale of EI into an observer-rated format (McCrae & Costa, 1987), which was rated by the peers.

Ability measure of Emotional Intelligence. Following the completion of the self-rated measures, participants completed the MSCEIT ability measure of emotional intelligence (Mayer, Salovey & Caruso, 2002). Mayer, Salovey, Caruso, and Sitarenios (2003) provide reliability and validity data for the test and its resulting factor structure. Given that Roberts et al. (2001) documented great similarity between the consensus scores generated by experts and lay people, we present scores based on the latter group in analyses. However, we also conducted all analyses below using expert scoring and found no differences in the results.

Supervisor-rated job performance. Supervisors completed a 5-item measure of general job performance from Côté and Miners (2006). Sample items include “How would you characterize your subordinate’s effectiveness in displaying job knowledge and skill?” and “How would you characterize your subordinate’s degree to which they set high standards and strive to meet them?”

Supervisor-rated transformational leadership ($\alpha=.88$). Supervisors completed the 22-item questionnaire on transformational leadership behaviors from Herold, Fedor, Caldwell, and Liu (2008). These authors adapted the items from the work of Rubin, Munz, and Bommer (2005) and Podsakoff, MacKenzie, and Bommer (1996). Sample items include “My subordinate provides individuals with new

ways of looking at things which are puzzling to them” and “My subordinate seeks new opportunities for our organization.” Although transformational leadership is frequently rated by subordinates rather than supervisors, subordinates did not take part in the students’ professional development exercise.

Supervisor-rated organizational citizenship behaviors (OCBs; $\alpha=.94$). Supervisors completed a 16-item scale of OCBs from Lee and Allen (2002). This instrument includes questions about the frequency of OCBs directed at helping individuals within the organization and at helping the organization as a whole. Sample items include “Willingly give his/her time to help others who have work-related problems” and “Show genuine concern and courtesy toward coworkers, even under the most trying business or personal situations.”

Demographic background. As in Study 1, participants’ self-reported sex, age, and whether English was their first language were used as control variables.

Personality scales. Participants completed self-report scales of Trait PA and Trait NA using the Positive And Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988). The ‘Big Five’ traits of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness were tested using the Ten Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003). Although the TIPI has relatively low internal consistency—given that the two-item scales for each trait were designed to emphasize content validity—it shows high convergent validity with other widely-used Big Five measures and very good test-retest reliability (mean $r = .72$ across traits; Gosling et al., 2003).

Cognitive intelligence. As a measure of cognitive intelligence, we used participants’ self-reported Graduate Management Admissions Test (GMAT) scores.

Results

Supplementary Table S3 contains bivariate correlations, descriptive statistics, and reliability values for study variables. The structure of the data in Study 2—as opposed to the data in Study 1—did not require the use of the Social Relations Model (Kenny & LaVoie, 1984) because each rater provided

responses about only one focal individual. With these statistically independent data, we were able to use conventional analyses to calculate the average values across all raters.

Consensus among judges

There was significant yet modest agreement across raters in making assessments of individuals' EI. Values for the intra-class correlation (ICC-1) representing the average agreement between any two raters were .071 for perceiving emotions, .028 for using emotions, .098 for understanding emotion, and .142 for managing emotions.

Self-observer agreement

As illustrated in Table 2, self-observer agreement was significant for all four branches, Perceiving ($r=.17$, 95% CI .00-.34), Use ($r=.19$, 95% CI .02-.36), Understanding ($r=.27$, 95% CI .10-.44), and Managing ($r=.28$, 95% CI .11-.45).

Convergent and divergent validity

All branches of EI except Use of Emotion met at least modestly the criteria of the multi-trait multi-model approach (MTMM; Campbell & Fiske, 1959), i.e., that the diagonal values in bold should be statistically significant and larger than any other value in the same row or column. This is the case for all factors except the Use of Emotion branch for one of the four values in the same column.

By contrast with self-reported EI, the ability test of EI showed no appreciable agreement with observer-reported EI. The results of a Confirmatory Factor Analysis of these three sources of data were consistent with this observation by visual inspection. A model converged but fit with only marginally acceptable statistics ($\chi^2=37.95$, $df=33$, $CFI=.98$, unrestricted log likelihood = -2035.01, $RMSEA=.03$, 95% CI .00-.08), with significant loadings onto the EI trait from self-rated and observer-rated EI (values for Perceiving, Use, Understanding, and Managing of .42/.51, .40/.49, .52/.61, and .85/.33, respectively, for self-rated/observer-rated EI). The only significant loading from the ability-tested factors of the MSCEIT was Managing (.25), with the remainder non-significant (-.16 to .08). There was substantial method

variance in all three sources: self-rated EI (.16-.69), observer-rated EI (-.14-.75), and ability-tested EI (.35-.73). More details of these analyses appear in Supplementary Table S4.

Predictive validity of observer ratings

Because these data were statistically independent, unlike those in Study 1 using the Social Relations Model, these results are based on conventional analyses rather than bootstrapping. The first 3 columns of Table 5 summarize the results of ordinary least squares multiple regression analyses in which observer-rated EI significantly predicts supervisor ratings, while controlling for personality, trait affect, cognitive ability and the other control variables. Each entry in the table is a regression coefficient from a separate model that includes age, gender, native English, Trait PA, Trait NA, Extraversion, Conscientiousness, Agreeableness, Neuroticism, and Openness. Baseline models including control variables appear in Supplementary Table S2. Because the observer closeness variable—even with its restricted range of 3 or higher—correlated highly with observer ratings of EI, we used this closeness rating as an additional control variable in these regression models. This helped to correct for the possible extraneous influence of interpersonal regard. Results show that higher observer-rated EI significantly predicts better supervisor ratings of transformational leadership, organizational citizenship, and job performance. Note that inclusion of cognitive intelligence as a control variable reduced the sample size substantially and also reduced the effect sizes for predictive validity, so that only transformational leadership remained significant. The first 3 columns of Supplementary Table S5 report coefficients for self-reported and ability-tested EI. Self-reported EI does not predict performance ratings, and ability-tested EI has unexpected negative associations.

Measurement equivalence of self and observer ratings

Table 4 presents the results of analyses of measurement equivalence between self and observer ratings of EI. The CFA models are described above in Study 1. The criterion was met for each nested model, in that the change in fit using the CFI index was smaller or equal to -0.01. This indicated support

for measurement equivalence.

Discussion

This study extends the body of evidence for understanding the social perception of emotional intelligence. First, the same basic pattern of results of the first study generalized to observer reports based on a different self-report instrument, which was developed to follow Mayer and Salovey's (1997) four-factor model of EI. Second, this study had greater external validity in that observer and supervisor ratings were collected from actual colleagues at participants' full-time jobs. The data included not only job performance, but also leadership and organizational citizenship behaviors, for which a link to EI has been theorized (Côté & Miners, 2006). Indeed, observer ratings of EI were positive predictors of all three criteria, above and beyond Big-Five personality factors, trait affectivity, and observer closeness to the targets—yet there were no such relationships apparent for self-rated EI. When controlling for cognitive intelligence in the subset of the sample for which data were available, transformational leadership, but not organizational citizenship or job performance ratings, remained significant.

Study 3

The two studies above, while providing worthwhile data regarding the social perception of EI, benefit from additional evidence to evaluate observer perceptions of emotional intelligence.

First, it is important to examine whether observer-reported EI is distinct from other observer-reported concepts, notably liking and positive regard, and other positive-valence judgments such as emotional stability that are potentially related to perceptions of EI. These socially desirable interpersonal constructs could be confounded with observer judgments of EI. With the goal of a comprehensive examination, we administered observer reports of liking, positive regard, and the big five personality traits. In addition to conducting structural equation models to test directly the divergence of observer-rated EI from these factors, we also used liking and positive regard as control variables in tests of

predictive validity.

Second, it is helpful to explore further the possible association between observer ratings and ability-tested EI. Study 2 used the MSCEIT test of EI (Mayer et al., 2002), and Study 3 uses the ability tests of EI developed by MacCann and Roberts (2008). Their situational judgment test of emotional understanding—that is, “understanding the relations between, and transitions among, emotions and between emotions and circumstances” (p. 540)—scores participant responses with respect to theoretical predictions based on the appraisal model of emotion (Ellsworth & Scherer, 2003; Lazarus, 1991; Roseman, 2001). Appraisal theory is a longstanding and well validated model of the process by which we interpret—even minimally—the stimuli in our environment in order to guide emotional experience. According to appraisal theory, there are a small number of dimensions by which we judge a situation in order to yield a specific emotional state. For example, a negative event can elicit emotions as diverse as anger, fear, sadness, and guilt—depending on whom we judge responsible for causing the event. The specific predictions of appraisal theory have been validated extensively across a range of cultural groups (e.g., Scherer, 1997). Likewise, MacCann and Roberts (2008) introduced a situational judgment test of emotion management. In this case, participants read vignettes regarding situations that could elicit strong emotions, and their responses are evaluated with respect to the modal responses of expert judges who were practitioners in professions that are geared towards emotional healing and relationship management. These tests have been used extensively in published research (e.g., Côté, DeCelles, McCarthy, Van Kleef, & Hideg, 2011).

Interestingly, the STEM has also been used in an observer-report format, with parents responding about their children (MacCann, Wang, Matthews, & Roberts, 2010). Instead of judging their children’s likely scores, as in other studies reporting peer-self agreement, participants provided predictions about the behaviors they believe their children would exhibit, and there was a correlation of $r=.19$ between the scores that resulted from these two sets of behavioral responses (MacCann, Wang, Matthews, & Roberts,

2010). This relatively modest overlap between self-rated hypothetical behaviors and the predictions made by closely knowledgeable observers suggests distinct perspectives between the two sources.

Method

Participants

Undergraduate students at a university on the East Coast of the US completed measures as part of their required introductory course in Management, and signed consent forms for additional research purposes. In this course, students were assigned into student teams averaging five members who worked closely throughout the semester. We collected observer ratings in the context of these teams. A total of $N=299$ individuals took part in this study, in $N=62$ teams. Because these measures were used to generate individualized assessments for developmental coaching as a required component of the course, there was missing data only for control variables, which led to $N=266$ for analyses involving controls.

Measures

Self-reported and Observer-reported EI. Respondents completed the same self-ratings and observer ratings of EI as in Study 2, adapted from Brackett et al. (2006). Each person completed ratings of each other team member, in a round robin format as in Study 1.

Ability tests of EI. The Situational Test of Emotional Understanding (STEU; MacCann & Roberts, 2008) contains 42 items describing situations, and participants select the likely emotions elicited by these situations using multiple-choice responses. Scoring was based on whether or not responses matched the predictions consistent with appraisal theory. The Situational Test of Emotion Management (STEM) contains 44 items describing situations, with multiple-choice responses indicating possible behavioral responses to these situations that are more vs. less appropriate. Scoring was based on the match of their responses with respect to expert ratings of appropriate behaviors in these situations.

Self-reported and Teammate reported Personality. Participants completed self-report scales of the Big

Five traits of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness using the Ten Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003). They also completed peer report version of the same scale for each teammate in a round robin format.

Observer ratings of Liking and Positive Regard. Participants rated their liking and positive regard for each teammate in a round robin format, on a scale from 1 (strongly disagree) to 7 (strongly agree). Liking used the questions: "I like my group member as a personal friend," and "I spend time (or would enjoy spending time) socially with my group member." Positive regard used the questions: "I have strong positive regard for my group member" and "I have deep respect for my group member." These two sets of questions correlated at $r=.93$, and so they appear together in the results below.

Performance ratings from teammates. For the purpose of developmental feedback, team members rated each other's performance. These ratings were collected at the end of the semester, with feedback aggregated to preserve the privacy of individual respondents. They rated each colleague's performance as an *Individual* and as a *Teammate*, using scales from 1 (not at all effective) to 7 (extremely effective).

Demographic background. As control variables, participants self-reported their sex, age, and whether English was their first language.

Trait affect. Participants completed measures of Trait PA and Trait NA using the Positive and Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988).

Cognitive intelligence. As a measure of cognitive intelligence, we used participants' self-reported Scholastic Achievement Test (SAT) scores.

Results

Supplementary Table S6 contains bivariate correlations, descriptive statistics, and reliability values for study variables.

Consensus among judges

Table 1 presents the results of Social Relations Model (SRM) analyses of emotional intelligence,

liking, and emotional stability. There was significant agreement about the EI of each particular target individual, both for individual branches as well as total EI (ranging from 16% to 23%). There was also significant target variance for liking (24%) and emotional stability (26%), which means that raters also tended to agree regarding who was more versus less likable and who was more versus less emotionally stable.

Note that there was moderate dyadic reciprocity in ratings of EI, $r=.20$ for Total EI, $p<.01$, which means that a rater who gave particularly high ratings to a target was, in turn, moderately likely to receive high ratings, from that target. For liking, however, dyadic reciprocity was substantially higher, $r=.44$, $p<.01$, which is consistent with past research showing that feelings of liking are particularly mutual (Kenny, 1994). This documents a divergence between the underlying properties of observer-rated EI and liking.

Self-observer agreement

As seen in Table 2, self-observer agreement was significant for the Perceiving ($r=.14$, 95% CI .02-.25), Use ($r=.27$, 95% CI .15-.38), and Managing ($r=.15$, 95% CI .03-.26) branches of EI, but not Understanding Emotion ($r=.11$, 95% CI -.01-.23).

Convergent and divergent validity

With the exception of the Understanding Emotion branch, the other three branches of EI demonstrate modest convergent and discriminant validity in terms of the criteria of the multi-trait multi-model approach (Campbell & Fiske, 1959). For these three factors, the diagonal values in bold are statistically significant and larger than underlined values in the same row or column. As in Studies 1 and 2, there was generally a positive manifold among the four branches of EI, for both self and observer ratings.

Degree of convergence between observer-rated and ability-tested EI

There was limited support for convergence between observer-reports and ability tests of EI. Peer

ratings of the Emotional Understanding branch correlated significantly with the STEU, $r=.12$, $p<.05$, particularly when controlling for peer ratings of Liking, partial $r=.20$, $p<.01$. Self-reported Emotional Understanding did not correlate with the STEU, $r=.08$, *ns*. However, this association appeared to lack divergent validity, in that the observer reports of emotional understanding have essentially the same association with ability-tested emotional understanding and emotion management ($r=.12$ and $r=.11$, respectively). Peer ratings of the Emotion Management branch did not correlate with the STEM, $r=-.04$, *ns*, even when controlling for Liking, $r=-.04$, *ns*. More details appear in Table 2.

Divergence of observer-rated EI from other observer-rated constructs

Table 6 summarizes confirmatory factor analyses (CFA) that support the notion that observer reports of EI are distinct from the related constructs of liking, positive regard, and emotional stability, i.e., low neuroticism. The latent factors were allowed to correlate, and these correlation matrices for all models are available upon request from the corresponding author. In each case, model fit was greatly improved by separating these constructs from EI, even while the overall model fit was fair. Also available upon request from the corresponding author are results for similar models that show the divergence of observer-reported EI from the traits of extraversion, agreeableness, openness, and conscientiousness.

Predictive validity of observer ratings

Columns 4-6 of Table 5 provided support for the predictive validity of observer-rated EI. The table presents the results of ordinary least squares multiple regression analyses in which observer-rated EI significantly predicts performance ratings from teammates. Each entry in the table is a regression coefficient from a separate model that includes age, gender, native English, and self-reports for the factors Trait PA, Trait NA, Extraversion, Conscientiousness, Agreeableness, Neuroticism, and Openness. Baseline models including control variables appear in Supplementary Table S2. Many coefficients for peer-rated EI are significant predictors of performance ratings whereas, by contrast, no coefficients were

significant for self-reported EI as reported in columns 4-6 of Supplementary Table S5. After controlling for Liking, the magnitude of the coefficients for observer-rated EI decreased in magnitude, but many remained significant—particularly for teammate versus individual effectiveness. Interestingly, the opposite was found for the STEM ability test of managing emotions—after controlling for liking, it predicted better individual but not teammate effectiveness. It is worth noting that the additional explanatory power of these variables was relatively modest. Note that, as mentioned above, these analyses exclude the two double-barreled questions from the Use of Emotion branch, and for comparison results are also presented with these two questions included, which tend to be negative predictors.

Measurement equivalence of self and observer ratings

Table 4 presents the results of analyses of measurement equivalence between self and observer ratings of EI, as described in more detail in Study 1. The criterion to conclude measurement equivalence was met for every stage of the model except the final constraint, i.e., the equality of mean values. This indicated evidence for leniency bias in this sample, where that self-reports were significantly higher than observer reports, but the underlying factor structure was otherwise equivalent.

Discussion

This study extended the body of evidence for evaluating the social perception of EI. First, it replicated the results of the first two studies demonstrating significant consensus among observers, significant self-observer agreement, and modest fit with criteria for construct validity. Next, there was modest and inconsistent evidence for the correspondence between observer-rated EI and performance on a respected ability test of EI. That is, observer-rated Emotion Understanding (EU) correlated with ability-tested EU, whereas self-reported EU did not—however, observer-reported management of emotion did not correlate with ability-tested emotion management, whereas self-reported emotion management did. Third, there was a significant distinction between observer-rated EI and the related interpersonal constructs of liking, positive regard, and emotional stability/neuroticism. Confirmatory factor analyses

showed better model fit when these constructs were distinct from observer-reported EI. In the case of liking, ratings of EI differed in terms of other underlying properties—notably, that there is large dyadic reciprocity for liking. Reciprocity is a key characteristic identified in the past literature for liking (Kenny, 1994), but it was only modest for observer-rated EI. Taken together, these empirical properties suggest a theoretical distinction between observer-rated EI and observer ratings of related constructs. Further, there was small yet significant predictive validity for observer-rated EI after controlling for liking and positive regard. This was the case for teammate effectiveness, but not individual effectiveness, which suggests that observer-rated EI is more relevant to interdependent tasks.

General Discussion

The current work provides the most comprehensive evidence to date about observer ratings of emotional intelligence (EI). We explored observer perceptions of emotional intelligence as epiphenomenal—that is, theoretically interesting because they exist inside our heads and have real consequences. Individuals make judgments of other people’s emotional abilities on a regular basis, regardless of whether doing so is useful or advisable. In attempting to understand these important judgments, three studies established a systematic body of evidence using authoritative methods from the study of interpersonal perception (Kenny & LaVoie, 1984) and construct validation (Campbell & Fiske, 1959), and replicating patterns of findings in three studies drawing from a total of 2,521 individuals.

In doing so, we mapped out a number of key psychometric properties of these social perceptions. First, observers tended at least modestly to agree with each other, and at least modestly with individuals’ own self ratings. Second, observer ratings tended to fit, also modestly, the construct validity criteria of the multi-trait multi-model (MTMM; Campbell & Fiske, 1959). Note that these MTMM analyses were a particularly conservative test, because they applied the criteria to distinguishing among the 4 factors of a four-factor model, rather than distinguishing observer-rated EI from other constructs. Third, there was consistent evidence for the predictive validity of observer-rated EI, above and beyond trait affect, the big

five personality traits, cognitive intelligence, age, gender, English as a native language, and closeness to the rater. Even when controlling for liking, observer-rated EI still predicted performance as a teammate, even if it no longer predicted performance as an individual contributor. Indeed, the better predictive power of observer-rated EI for team vs. individual performance suggests that the construct is tapping into a form of effectiveness that is interpersonal—rather than intrapersonal. Although this study focused specifically on workplace-relevant criteria in establishing predictive validity, this result highlights the importance of EI in the realm of interpersonal behavior more generally. Fourth, there was good evidence that observer-rated EI is distinct from related judgments that could be considered possible confounding factors, notably liking and emotional stability. In addition to this direct evidence for divergent validity in the form of confirmatory factor analysis, the MTMM findings across the four branches of EI also provided further evidence for divergent validity. That is, if the four factors were distinct from each other, then at least some branches must also be distinct from these confounding factors. If judges were merely guided by liking or even stereotypes about the person high in EI, then these nuisance factors should apply equally across the branches of EI. Further, observer-rated EI had different statistical properties from liking—notably, we replicated the longstanding finding that liking is highly mutual (Kenny, 1994), but much less so for ratings of EI. Last, somewhat surprisingly there was only poor evidence that observer ratings of EI correlate with EI ability tests. We discuss the possible interpretation and meaning of this below.

Are observer perceptions of EI valid?

Although we explored observer-rated emotional intelligence as an epiphenomenal construct—where these ratings are important regardless of whether they reflect an underlying reality—it is also necessary to pose the question of whether observers can actually detect another person’s EI. The evidence in this paper suggests caution in making this conclusion. First, many positive results that were statistically significant had modest effect sizes, and some confirmatory factor models showed that fit

increased with nested models and yet fit was poor overall. Importantly, there was no clear evidence for any convergence between observer-rated EI and ability-tested EI. One would expect to see this, to the extent that existing ability tests have been validated as representing the EI construct. That said, it is worth noting that studies using multiple ability tests of EI typically find that these tests do not correlate strongly with each other—which Mayer et al. (2008) in their *Annual Review of Psychology* chapter of emotional intelligence described as “both perplexing and troubling” (p. 518). As such, one would not necessarily require strong associations between observer reports and ability tests—even so, the lack of any consistent association is concerning. It is worthwhile to explore whether there may be moderating conditions under which EI ratings are valid, and potential methods to collect these judgments that could maximize their validity. Notably, raters who are closer would have a greater basis of experience from which to draw in making their judgments. Further, those who are themselves high in emotional abilities are likely to be more capable of judging other people’s emotional abilities—a ‘takes one to know one’ type of effect (Carney & Harrigan, 2003).

Limitations and Future Research

A number of limitations qualify the results of this research, and suggest fertile ground for additional work.

First, the present study adapted self-report questionnaires for use in an observer-rating format, in keeping with norms from the field of personality (McCrae & Costa, 1987). However, in the domain of emotional intelligence some areas may be too intra-psychic for observers to be able to report validly. Particularly in Studies 2 and 3, we used a self-report scale that was developed to match the content domain of an emotional intelligence ability test (MSCEIT, Mayer, Salovey & Caruso, 2002) —which has limited content in that some areas of EI may not have been emphasized sufficiently if they were not amenable to ability testing. We suggest that future research on observer-rated EI should not start with a self-report instrument for adaptation, but rather should start with the theoretical question of which

emotional skills should leave visible behavioral traces that observers might be able to see.

Second, these studies examined emotional intelligence only in the context of workplace settings, and EI is crucial for interpersonal interaction in relationships more generally. The primary reason for the focus on the workplace is that most adults spend the majority of their waking hours at work, and for this reason it seemed like an important context to examine. Further, the participants and observers we sampled shared strong acquaintance with each other and had meaningful outcomes to their interactions together. Future work could examine social groups with other types of personal relationships, such as romantic partners, families, dormitories, fraternities and sororities, houses of worship, and military organizations.

Finally, the research presented here was observational rather than experimental. As such, we did not have the opportunity to observe the detailed psychological processes that contribute to observer judgments of EI. Having access to the particular cues that are given off by individuals higher vs. lower in EI, as well as the cues utilized by perceivers, would allow a lens model analysis (Brunswick, 1955) to better understand the social perception process. Notably, some cues may be used successfully, whereas others might involve inaccurate stereotypes or red herrings.

The use of observer ratings: Promising but no panacea

Observer ratings of EI could be seen as having a potential practical application in academic research. Measurement challenges in the field have long been noted, and observer ratings might be seen as a solution to that challenge. It is worthwhile to evaluate this potential choice in the context of the four criteria that Matthews, Zeidner, and Roberts (2002) outlined in their large-scale review and critique of research on emotional intelligence, which any purported measure of EI should at least minimally satisfy: (a) content validity, (b) scale reliability, (c) construct validity, including convergent and divergent validity, and (d) predictive validity for relevant criteria. Observer reports were found here to satisfy these four requirements to varying extents. As such, using observer reports to measure EI must be done with

caution. We note that these criteria are satisfied with statistical significance yet often with modest effect sizes, modest model fit, and modest variance explained in predictive models. Along these lines, although there was significant consensus across judges, a relatively large number of judges would be required to create a composite score that approaches conventional levels of total reliability. Based on the effect sizes reported above, it would require at least 12 raters to achieve total reliability in the .90 range for each component of EI, and at least 4 raters for a .70 level of reliability. That said, one of the benefits of observer ratings is that one can aggregate information from as many informants as needed—whereas, by contrast, there is only one self who is available for self-ratings (Paulhus et al., 1998). In any use of observer reports, each rater should ideally rate multiple targets, so that any individual-level rater bias in how people use the response scales can be measured and corrected. To the extent that the results above provide only modest evidence for the validity of observer-reported EI, based on what is known at this point, we suggest that their use be limited to research and educational contexts, rather than evaluative domains. The data in this study were collected in classroom contexts that emphasized the value of direct and honest feedback for the sake of colleagues' learning, and we do not know how results would generalize to settings that have instrumental stakes for the raters and targets. Future research can enhance our knowledge of when and how observer ratings of EI might be useful as a measurement device.

This paper argues that we judge each other's social skills and we use these judgments—for better or for worse. In the modern workplace, such judgments are often explicit hiring criteria, and they are typically made by interviewers on an idiosyncratic rather than systematic basis. In social settings, people use these judgments to steer towards and away from potential relationships. Along these lines, the open question about the validity of social perceptions does not mean they are not prevalent, important, and potentially useful. In the case of coaching for professional development and leadership training, it is valuable to know what other people think of you—given that other people are using these impressions.

The strongest findings in this paper were the consistent results for predictive validity—we found in each study that observer ratings of EI predict performance ratings from both supervisors and peers. Thus individuals can likely benefit from feedback about others’ perceptions of them, whether or not they find the feedback accurate. We conclude that observer judgments of EI are consequential. Taken together, this study offers a substantial body of evidence that better helps us to understand the nature of observer judgments of emotional intelligence. These judgments are an important phenomenon for researchers to understand—because they are formed and used to navigate interpersonal relationships, regardless of what they might mean.

References

- Averill, J. R. (2004). A tale of two snarks: Emotional intelligence and emotional creativity compared. *Psychological Inquiry*, 3, 228-233.
- Bar-On, R., & Handley, R. (2003). *The BarOn EQ-360: Technical manual*. Toronto, Canada: Multi-Health Systems.
- Boyatzis, R. E., & Sala, F. (2004). The Emotional Competence Inventory (ECI). In G. Geher (Ed.), *Measuring Emotional Intelligence* (pp. 147-180). Nova Science.
- 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, 780-795.
- Brackett, M., Mayer, J. D. & Warner, R.M. (2004). Emotional intelligence and its relation to everyday behaviour. *Personality and Individual Differences*, 36, 1387-1402.
- Brunswick, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193-217.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Carmeli, A., & Josman, Z. E. (2006). The relationship among emotional intelligence, task performance, and organizational citizenship behaviors. *Human Performance*, 19, 403-419.
- Carney, D. R., & Harrigan, J. A. (2003). It takes one to know one: Interpersonal sensitivity is related to accurate assessments of others' interpersonal sensitivity. *Emotion*, 3, 194-200.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Charbonneau, D., & Nicol, A. A. M. (2002). Emotional intelligence and leadership in adolescents. *Personality and Individual Differences*, 33, 1101-1113.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233-255.
- Costa, Jr., P. T. & McCrae, R. R. (1992). Normal Personality Assessment in Clinical Practice: The NEO Personality Inventory. *Psychological Assessment*, 4, 5-13.

- Costanzo, M., & Archer, D. (1989). Interpreting the expressive behavior of others: The Interpersonal Perception Task. *Journal of Nonverbal Behavior, 13*, 225-245.
- Côté, S. (1999). Affect and performance in organizational settings. *Current Directions in Psychological Science, 8*, 65-68.
- Côté, S., DeCelles, K. A., McCarthy, J. M., Van Kleef, G. A., & Hideg, I. (2011). The Jekyll and Hyde of Emotional Intelligence Emotion-Regulation Knowledge Facilitates Both Prosocial and Interpersonally Deviant Behavior. *Psychological Science, 22*, 1073-1080.
- Côté, S., & Miners, C. T. H. (2006). Emotional intelligence, cognitive intelligence, and job performance. *Administrative Science Quarterly, 51*, 2006, 1-28
- D'Augelli, A. R. (1973). The assessment of interpersonal skills: A comparison of observer, peer, and self ratings. *Journal of Community Psychology, 1*, 177-179.
- Davies, M., Stankov, L., & Roberts, R. D. (1998). Emotional intelligence: In search of an elusive construct. *Journal of Personality and Social Psychology, 75*, 989-1015.
- Day, A. L., & Carroll, S. A. (2004). Using an ability-based measure of emotional intelligence to predict individual performance, group performance, and group citizenship behaviours. *Personality and Individual Differences, 36*, 1443-1458.
- Downs, A., Strand, P & Cerna, S. (2007). Emotion understanding in English- and Spanish-speaking preschoolers enrolled in HeadStart. *Social Development, 16*, 410-439.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics, 7*, 1-26.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
- Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal processes in emotion. In R. J. Davidson, H. Goldsmith & K. R. Scherer (Eds.), *Handbook of the Affective Sciences* (pp. 572-595). New York and Oxford: Oxford University Press.
- Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science, 2*, 31-39.
- Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55*, 149-158.

Funder, D. C., & Colvin, C. R. (1997). Congruence of others' and self-judgments of personality. In Hogan, Robert (Ed); Johnson, John A. (Ed); et al. *Handbook of personality psychology*. (pp. 617-647). San Diego: Academic Press.

Funder, D. C., & West, S. G. (1993). Consensus, self-other agreement, and accuracy in personality judgment: An introduction. *Journal of Personality*, *61*, 457-476.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A Very Brief Measure of the Big Five Personality Domains. *Journal of Research in Personality*, *37*, 504-528.

Gutman, L. & Levy, S. (1991). Two structural laws for intelligence. *Intelligence*, *15*, 79-103.

Hammond, K. R. & Stewart, T. R. (Eds.) (2001). *The essential Brunswik: Beginnings, explications, applications*. London, England: Oxford University Press.

Harms, P. D., & Credé, M. (2010). Emotional Intelligence and Transformational and Transactional Leadership: A meta-analysis. *Journal of Leadership and Organizational Studies*, *17*, 5-17.

Harris, M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, *41*, 43-62.

Herold, D. M., Fedor, D. B., Caldwell, S., & Liu, Y. (2008). The effects of transformational and change leadership on employees' commitment to a change: A multilevel study. *Journal of Applied Psychology*, *93*, 346-357.

James, L. R., Demaree, R. G., & Wolf, G. (1993). r_{wg} : An assessment of within-group interrater agreement. *Journal of Applied Psychology*, *78*, 306-309.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, *61*, 521-551.

John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102-138). New York: Guilford Press.

Jordan, P. J., & Ashkanasy, N. M. (2005). Emotional intelligence, emotional self-awareness, and team effectiveness. In V. U. Druskat, F. Sala, & G. J. Mount (Eds.), *The impact of emotional intelligence on individual and group performance* (pp. 145-163). Mahwah, NJ: Lawrence Erlbaum Associates.

Joseph, D. L., Jin, J., Newman, D. A., & O'Boyle, E. H. (2014). Why does self-reported emotional intelligence predict job performance? A meta-analytic investigation of mixed EI. *Journal of Applied Psychology*, Provisionally accepted.

Joseph, D. L., & Newman, D. A. (2010). Discriminant validity of self-reported emotional intelligence: A multitrait-multisource study. *Educational and Psychological Measurement*, 70, 672-694.

Kellett, J. B., Humphrey, R. H., & Sleeth, R. G. (2006). Empathy and the emergence of task and relations leaders. *Leadership Quarterly*, 17, 146-162.

Kenny, D. A. & LaVoie, L. (1984). The Social Relations Model. *Advances in Experimental Social Psychology*, 18, 141-182.

Kenny, D. A. (1994). *Interpersonal Perception: A Social Relations Analysis*. New York: Guilford Press.

Kenny, D.A. (1998). SOREMO [computer program]. <http://nw3.nai.net/~dakenny/kenny.htm>

Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality*, 64, 311-337.

Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research*, 75, 63-82.

Law, K. S., Wong, C., & Song, L. J. (2004). The Construct and Criterion Validity of Emotional Intelligence and Its Potential Utility for Management Studies. *Journal of Applied Psychology*, 89, 483-496.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Lee, K., & Allen, N. J. (2002). Organizational citizenship behavior and workplace deviance: The role of affect and cognitions. *Journal of Applied Psychology*, 87, 131-142.

Libbrecht, N., Lievens, F., & Schollaert, E. (2010). Measurement equivalence of the Wong and Law Emotional Intelligence Scale across self and other ratings. *Educational and Psychological Measurement*, 70, 1007-1020.

Lopes, P. N., Grewal, D., Kadis, J., Gall, M., & Salovey, P. (2006). Evidence that emotional intelligence is related to job performance and affect and attitudes at work. *Psicothema*, 18, 132-138.

Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*, 280-296.

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

MacCann, C., Wang, L., Matthews, G., & Roberts, R. D. (2010). Emotional intelligence and the eye of the beholder: Comparing self-and parent-rated situational judgments in adolescents. *Journal of Research in Personality, 44*, 673-676.

Matthews, G., Roberts, R. D., & Zeidner, M. (2004). Seven myths about emotional intelligence. *Psychological Inquiry, 15*, 179-196.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence: Science and Myth*. Cambridge, MA: MIT Press.

Mayer, J. D. & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp.3-31). New York: Basic Books.

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 27*, 267-298.

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human Abilities: Emotional intelligence. *Annual Review of Psychology, 58*, 507-536.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional Intelligence: New Ability or Eclectic Traits? *American Psychologist, 63*, 503-517.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2000). Models of emotional intelligence. In R.J. Sternberg (ed), *Handbook of Intelligence*, pp.396-420. Cambridge, UK: Cambridge University Press.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2002). *Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) User's Manual*. Toronto, Canada: MHS Publishers.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring Emotional Intelligence with the MSCEIT V2.0. *Emotion, 3*, 97-105.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81-90.

- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*, 246-268.
- Modassir, A., & Singh, T. (2008). Relationship of emotional intelligence with transformational leadership and organizational citizenship behavior. *International Journal of Leadership Studies*, *4*, 3-21.
- Newman, D. A., & Lyon, J. S. (2009). Recruitment efforts to reduce adverse impact: Targeted recruiting for personality, cognitive ability, and diversity. *Journal of Applied Psychology*, *94*, 298-317.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.
- O'Boyle, E. H., Humphrey, R., Pollack, J., Hawver, T. H., & Story, P. A. (2011). Emotional intelligence and job performance: A meta-analysis. *Journal of Organizational Behavior*, *32*, 788-818.
- Oatley, K. (2004). Emotional intelligence and the intelligence of emotions. *Psychological Inquiry*, *15*, 216-221.
- Paulhus, D. L., Lysy, D. C., & Yik, M. S. M. (1998). Self-report measures of intelligence: Are they useful as proxy IQ tests? *Journal of Personality*, *66*, 525-554.
- Paulhus, D., & Morgan, K. L. (1997). Determinants of perceived intelligence in leaderless groups: The dynamic effects of shyness and familiarity. *Journal of Personality and Social Psychology*, *72*, 99-107.
- Petrides, K.V., Furnham, A., & Mavroveli, S. (2007). Trait emotional intelligence: Moving forward in the field of EI. In Matthews G, Zeidner M, Roberts R. (Eds). *The science of emotional intelligence: Knowns and unknowns* (pp. 151-166). New York: Oxford University Press.
- Podsakoff, P. M., MacKenzie, S. B., & Bommer, W. (1996). A meta-analysis of the relationships between Kerr and Jermier's substitutes for leadership and employee job attitudes, role perceptions, and performance. *Journal of Applied Psychology*, *81*, 380-399.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879-903.
- Redzanowski, U., & Glück, J. (2013). Who knows who is wise? Self and peer ratings of wisdom. *Journal of Gerontology, Series B, Psychological Sciences and Social Sciences*, *68*, 391-394.

Rivers, S.E., Brackett, M.A., Salovey, P., & Mayer, J.D. (2007). Measuring emotional intelligence as a set of mental abilities. In G. Matthews, M. Zeidner, & R.D. Roberts (Eds.), *The science of emotional intelligence: Knowns and unknowns* (pp. 230-257). New York: Oxford University Press.

Roberts, R. D., Zeidner, M., & Matthews, G. (2001). Does emotional intelligence meet traditional standards for an intelligence? Some new data and conclusions. *Emotion, 1*, 196-231.

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

Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, P. L., & Archer, D. (1979). *Sensitivity to nonverbal communication: The PONS Test*. Baltimore: Johns Hopkins University Press.

Rubin, R. S., Munz, D. C., & Bommer, W. H. (2005). Leading from within: The effects of emotion recognition and personality on transformational leadership behavior. *Academy of Management Journal, 48*, 845-858.

Sala, F. (2005). *Emotional Competence Inventory: Technical Manual*. Hay Group. Retrieved February 26, 2009, from <http://www.eiconsortium.org/>.

Scherer, K. (1997). The role of culture in emotion-antecedent appraisal. *Journal of Personality and Social Psychology, 73*, 902-922.

Song, L. J., Huang, G. H., Peng, K. Z., Law, K. S., Wong, C. S., & Chen, Z. J. (2010). The differential effects of general mental ability and emotional intelligence on academic performance and social interactions. *Intelligence, 38*, 137-143.

Staw, B. M., & Barsade, S. G. (1993). Affect and managerial performance: A test of the sadder-but-wiser vs. happier-and-smarter hypotheses. *Administrative Science Quarterly, 38*, 304-331.

Staw, B.M., Sutton, R.I., & Pelled, L.H. (1994). Employee positive emotion and favorable outcomes at the workplace. *Organization Science, 5*, 51-71.

Swann, W. B., Jr. Chang-Schneider, C. & McClarty, K. (2007). Do people's self-views matter? Self-concept and self-esteem in everyday life. *American Psychologist, 62*, 84-94.

Swann, W. B., Jr., Stein-Seroussi, A., & McNulty, S. E. (1992). Outcasts in a white lie society: The enigmatic worlds of people with negative self-conceptions. *Journal of Personality and Social Psychology, 62*, 618-624.

Tellegen, A. (1982). *Brief manual of the differential personality questionnaire*. Minneapolis: University of Minnesota Press.

Tett, R. P., Fox, K. E., and Wang, A. (2005). Development and Validation of a Self-Report Measure of Emotional Intelligence as a Multidimensional Trait Domain. *Personality and Social Psychology Bulletin*, 31, 859-888.

Van der Zee, K., Thijs, M., & Schakel, L. (2002). The relationship of emotional intelligence with academic intelligence and the Big Five. *European Journal of Personality*, 16, 103-125.

Watson, D., & Clark, L. A. (1991). Self- versus peer ratings of specific emotional traits: Evidence of convergent and discriminant validity. *Journal of Personality and Social Psychology*, 60, 927-940.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.

Wechsler, D. (1944). *The measurement of adult intelligence (3rd ed.)*. Baltimore: Williams and Wilkins.

Wong, C. S., & Law, K. S. (2002). The effects of leader and follower emotional intelligence on performance and attitude: An exploratory study. *Leadership Quarterly*, 13, 243-274.

Table 1
 Variance partitioning of observer-rated Emotional Intelligence (Studies 1 and 3)

Factor	Source of Variance			
	Target	Rater	Relationship/Error	Dyadic Reciprocity
<i>a. Study 1</i>				
<i>Emotional Intelligence</i>				
Self-emotions appraisal	9% **	40% **	51%	0.02
Other-emotions appraisal	21% **	28% **	51%	0.14 **
Use of emotion	25% **	26% **	49%	0.09 **
Regulation of emotion	21% **	28% **	50%	0.05
Total EI	18% **	41% **	41%	0.13 **
<i>b. Study 3</i>				
<i>Emotional Intelligence</i>				
Perceiving Emotions	16% ***	26% ***	58%	0.16
Use of Emotions	21% ***	12% ***	67%	-0.11 *
Understanding Emotions	11% ***	34% ***	55%	0.15 **
Managing Emotions	23% ***	24% ***	54%	0.19
Total EI	21% ***	26% ***	53%	0.20 **
<i>Liking/Regard</i>	24% ***	26% ***	50%	0.44 ***
<i>Emotional stability</i>	26% ***	17% ***	57%	-0.01

* $p < .05$; ** $p < .01$; all values two-tailed.

Notes: Study 1 included 3,990 ratings from $N=931$ individuals in 186 teams. Study 3 included 1,153 ratings from 299 individuals in 62 teams. Boldface indicates coefficients representing consensus among judges.

Table 2
Multi-trait multi-method analysis of Emotional Intelligence ratings (Studies 1, 2, and 3)

Variable	α	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
<i>a. Study 1, (N=931)</i>															
<i>Self-rated Emotional Intelligence</i>															
1. Self-emotions appraisal	.84														
2. Other-emotions appraisal	.86	.42 ***													
3. Use of emotion	.76	.33 ***	.20 ***												
4. Regulation of emotion	.90	.36 ***	.32 ***	.34 ***											
5. Total Self-rated EI	.87	.71 ***	.69 ***	.61 ***	.78 ***										
<i>Observer-rated Emotional Intelligence</i>															
6. Self-emotions appraisal	.92	.08 *	.08 *	.05	.09 **	.11 ***									
7. Other-emotions appraisal	.93	.06	.14 ***	.00	.12 ***	.12 ***	.65 ***								
8. Use of emotion	.84	.02	-.06	.24 ***	.03	.06 *	.37 ***	.29 ***							
9. Regulation of emotion	.93	.05	.06	-.02	.25 ***	.14 ***	.48 ***	.59 ***	.19 ***						
10. Total Observer-rated EI	.93	.06 *	.08 *	.08 *	.17 ***	.15 ***	.79 ***	.86 ***	.59 ***	.78 ***					
<i>b. Study 2 (N=135)</i>															
<i>Self-rated Emotional Intelligence</i>															
1. Perceiving emotion	.74														
2. Use of emotion	.87	.22 *													
3. Understanding emotion	.87	.37 **	.30 **												
4. Managing	.67	.45 **	.01	.25 **											
5. Total Self-rated EI	.85	.73 **	.38 **	.75 **	.74 **										
<i>Observer-rated Emotional Intelligence</i>															
6. Perceiving emotion	.78	.17 *	.14	.05	.09	.15									
7. Use of emotion	.57	.07	.19 *	.14	-.14	.06	-.01								
8. Understanding emotion	.82	.09	.20 *	.27 **	.06	.21 *	.50 **	.21 *							
9. Managing	.74	.04	.07	-.03	.28 **	.15	.64 **	-.27 **	.33 **						
10. Total Observer-rated EI	.86	.13	.19 *	.11	.19 *	.22 *	.83 **	.08	.70 **	.85 **					
<i>Ability-tested Emotional Intelligence</i>															
11. Perceiving emotion	-	.06	.00	.14	.03	.09	-.07	.08	.07	-.11	-.04				
12. Use of emotion	-	.11	-.01	.00	.12	.08	-.02	.00	.04	-.03	.00	.61 **			
13. Understanding emotion	-	.02	-.10	.15	-.02	.04	.02	-.01	.16	-.01	.05	.28 **	.26 **		
14. Managing	-	.27 **	.04	.13	.30 **	.28 **	.10	.17	.16	.04	.14	.25 **	.34 **	.27 **	
15. Total Peer-rated EI	-	.15	.00	.15	.13	.16	-.01	.08	.14	-.05	.04	.84 **	.80 **	.55 **	.57 **
<i>c. Study 3 (N=299)</i>															
<i>Self-rated Emotional Intelligence</i>															
1. Perceiving emotion	.72														
2. Use of emotion	.75	.04													
3. Understanding emotion	.94	.37 ***	.23 ***												
4. Managing	.75	.52 ***	-.08	.27 ***											
5. Total Self-rated EI	.82	.71 ***	.36 ***	.74 ***	.75 ***										
<i>Observer-rated Emotional Intelligence</i>															
6. Perceiving emotion	.83	.14 *	.13 *	.11 *	.12 *	.19 **									
7. Use of emotion	.67	-.03	.27 ***	.06	-.05	.07	.12 *								
8. Understanding emotion	.89	.02	.17 **	.11	.06	.13 *	.68 ***	.21 ***							
9. Managing	.87	.07	-.04	-.03	.15 **	.07	.74 ***	-.09	.52 ***						
10. Total Observer-rated EI	.90	.07	.11	.06	.12 *	.14 *	.89 ***	.24 ***	.79 ***	.89 ***					
<i>Ability-tested Emotional Intelligence</i>															
11. Understanding emotion	.69	.03	.06	.08	.07	.10	-.04	-.03	.12 *	-.06	-.02				
12. Managing	.85	-.03	.03	.10	.13 *	.11	-.04	.08	.11	-.04	.01	.59 ***			

*p < .05; **p < .01; all values two-tailed.

Notes: Bold typeface refers to "monotrait-heteromethod" coefficients, i.e., peer-self agreement. Italicized typeface refers to "monomethod-heterotrait" coefficients, i.e., convergence among branches. Underlined typeface refers to "heterotrait-heteromethod" coefficients, i.e., challenges to divergent validity across branches.

Table 3

Partial correlations between Emotional Intelligence and Performance Ratings, Controlling for Demographic Variables, Personality Traits, and Cognitive Intelligence (Study 1)

Emotional Intelligence (EI)	Performance Ratings from Teammates					
	Individual		Teammate		Leader	
	M	95% CI	M	95% CI	M	95% CI
	<i>a. Observer-rated EI</i>					
Self-emotions appraisal	.09 **	(.03, .16)	.11 **	(.04, .17)	.11 **	(.05, .17)
Other-emotions appraisal	.12 **	(.06, .19)	.16 **	(.10, .22)	.14 **	(.08, .21)
Use of emotion	.25 **	(.19, .31)	.23 **	(.17, .29)	.17 **	(.11, .23)
Regulation of emotion	.08 **	(.01, .15)	.11 **	(.04, .17)	.10 **	(.04, .17)
Total EI	.16 **	(.10, .23)	.19 **	(.12, .25)	.16 **	(.10, .23)
	<i>b. Self-rated EI</i>					
Self-emotions appraisal	.04	(-.04, .11)	.04	(-.03, .11)	.00	(-.06, .06)
Other-emotions appraisal	.02	(-.07, .10)	.05	(-.03, .13)	.04	(-.03, .11)
Use of emotion	.16 **	(.09, .23)	.13 **	(.05, .20)	.08 *	(.01, .15)
Regulation of emotion	-.02	(-.09, .05)	.02	(-.06, .09)	-.04	(-.11, .04)
Total EI	.06	(-.02, .14)	.08 *	(.00, .16)	.03	(-.04, .10)

* $p < .05$; ** $p < .01$; all values two-tailed.

Notes: $N=931$ individuals in 186 teams. Estimates for peer-rated EI are based on bootstrap simulations that sample one peer for performance ratings and a distinct peer for EI ratings, in order to reduce same-source bias. The estimates and significance levels of the bootstrap simulations are based on 10,000 samples. These partial correlations control for age, gender, native english speaker, positive affect, negative affect, extraversion, conscientiousness, agreeableness, neuroticism, openness to experience, and total GMAT scores.

Table 4

Analysis of Measurement Equivalence between Observer-reported and Self-reported Emotional Intelligence (Studies 1, 2, and 3)

<i>Confirmatory Factor Analysis</i>	<i>Chi-square</i>	<i>df</i>	<i>CFI</i>	<i>Change</i>		
				<i>in CFI</i>	<i>RMSEA</i>	<i>BIC</i>
<i>a. Study 1</i>						
<i>Model 1: Configural Invariance</i>	2428.0	196	.972	-	.060	227472.3
<i>Model 2: Weak Invariance (equal loadings)</i>	2754.1	208	.969	.004	.063	227693.4
<i>Model 3: Strong Invariance (equal loadings and intercepts)</i>	3029.7	220	.965	.003	.064	227864.1
<i>Model 4: Equal loadings, intercepts, and means</i>	3248.0	224	.963	.003	.066	228047.4
<i>b. Study 2</i>						
<i>Model 1: Configural Invariance</i>	538.6	196	.954	-	.057	60531.3
<i>Model 2: Weak Invariance (equal loadings)</i>	561.9	208	.952	.002	.056	60470.7
<i>Model 3: Strong Invariance (equal loadings and intercepts)</i>	632.8	220	.944	.008	.059	60457.7
<i>Model 4: Equal loadings, intercepts, and means</i>	717.4	224	.933	.011	.064	60514.4
<i>c. Study 3</i>						
<i>Model 1: Configural Invariance</i>	1213.7	196	.918	-	.086	78108.7
<i>Model 2: Weak Invariance (equal loadings)</i>	1267.9	208	.914	.003	.085	78075.8
<i>Model 3: Strong Invariance (equal loadings and intercepts)</i>	1395.6	220	.905	.009	.087	78116.6
<i>Model 4: Equal loadings, intercepts, and means</i>	1606.4	224	.888	.017	.094	78298.3

Table 5

Regression coefficients illustrating predictive validity for observer-rated Emotional Intelligence, Controlling for Demographic Variables, Personality Traits, and Cognitive Intelligence (Studies 2 and 3)

	Study 2			Study 3			
	Performance Ratings from Supervisors			Performance Ratings from Teammates		Performance Ratings from Teammates Controlling for Liking	
	Organizational Citizenship	Transformational Leadership	Job Performance	Individual Effectiveness	Individual Teammate	Individual Effectiveness	Teammate Effectiveness
a. Without controlling for Cognitive Intelligence							
	N=135			N=283			
Perceiving emotion	.16	.27 ** (.06)	.09	.49 *** (.21)	.65 *** (.40)	.00	.17 *** (.01)
Use of emotion	.22 * (.04)	.21 * (.04)	.12	.22 *** (.04)	.30 *** (.08)	-.07	-.03
Understanding emotion	.11	.11	.07	.50 *** (.21)	.57 *** (.27)	.17 ** (.02)	.18 *** (.02)
Managing emotion total	.26 ** (.05)	.34 *** (.09)	.17	.45 *** (.20)	.66 *** (.42)	-.05	.19 *** (.02)
Total EI	.26 **	.34 *** (.09)	.16	.52 *** (.25)	.70 *** (.47)	.02	.24 *** (.02)
CI for total EI	(.07, .45)	(.16, .53)	(-.03, .36)	(.42, .62)	(.61, .79)	(-.11, .16)	(.13, .34)
df	(1, 125)	(1, 125)	(1, 125)	(1, 272)	(1, 272)	(1, 271)	(1, 271)
b. Controlling for Cognitive Intelligence							
	N=97			N=266			
Perceiving emotion	.16	.27 ** (.09)	.08	.52 *** (.25)	.68 *** (.43)	.06	.20 *** (.02)
Use of emotion	.17	.18	.15	.24 *** (.05)	.29 *** (.08)	-.05	-.04
Understanding emotion	.08	.12	.08	.50 *** (.21)	.57 *** (.29)	.17 ** (.02)	.18 *** (.02)
Managing emotion total	.19	.26 ** (.06)	.09	.48 *** (.22)	.66 *** (.44)	.00	.21 *** (.02)
Total EI	.18	.26 ** (.07)	.09	.54 *** (.28)	.71 *** (.49)	.08	.26 *** (.03)
CI for total EI	(.03, .39)	(.06, .46)	(-.09, .28)	(.44, .64)	(.62, .79)	(-.06, .21)	(.15, .36)
df	(1, 86)	(1, 86)	(1, 86)	(1, 254)	(1, 254)	(1, 253)	(1, 253)

* $p < .05$; ** $p < .01$; all values two-tailed.

Notes: For significant coefficients, the incremental R-squared appears in parentheses Double-barreled questions in the Use of Emotion branch are removed from analysis, but appear in Supplementary Table 6.

Table 6

Confirmatory Factor Analysis models distinguishing Observer-rated Emotional Intelligence from Related Constructs (Study 3)

<i>Fit statistic</i>	<i>Distinguishing EI and Liking</i>			<i>Distinguishing EI and Observer-rated Neuroticism</i>		
	<i>Global Factor</i>	<i>Two-Factor Model</i>	<i>Five-Factor Model</i>	<i>Global Factor</i>	<i>Two-Factor Model</i>	<i>Five-Factor Model</i>
	<i>Combining EI and Liking</i>	<i>Separating EI and Liking</i>	<i>Separating EI Four Factors and Liking</i>	<i>Combining EI and Neuroticism</i>	<i>Separating EI and Neuroticism</i>	<i>Separating EI Four Factors and Neuroticism</i>
Chi-Square	2315.7 **	1814.6 **	1300.4 **	1866.4 **	1714.2 **	1132.5 **
df	252	251	242	189	188	179
Chi-Square/df	9.19	7.23	5.37	9.88	9.12	6.33
Baseline Chi-Square	5939.9 **	5939.9 **	5939.9 **	4405.7 **	4405.7 **	4405.7 **
Change in Chi-square	-	501.2 **	514.2 **	-	152.2 **	581.7 **
CFI	.64	.72	.81	.60	.64	.77
TFI	.60	.70	.79	.56	.59	.73
Log-Likelihood	-8358.2	-8107.6	-7850.5	-7629.4	-7553.3	-7262.4
Unrestricted Log-Likelihood	-7200.4	-7200.4	-7200.4	-6696.2	-6696.2	-6696.2
Number of Parameters	48	49	58	42	43	52
AIC	16812.4	16313.3	15817.1	15342.7	15192.5	14628.8
BIC	16990.1	16494.6	16031.7	15498.1	15351.6	14821.2
BIC2	16837.8	16339.2	15847.8	15365.0	15215.3	14656.3
RMSEA	.17** (.16-.17)	.14** (.14-.15)	.12** (.11-.13)	.17** (.17-.18)	.16** (.16-.17)	.13** (.13-.14)

* $p < .05$; ** $p < .01$; all values two-tailed.

Notes: N=299. Numbers in parentheses indicate lower and upper bounds, respectively. Degrees of freedom for the baseline model are 276 for analyses of EI and Liking, and 210 for EI and Neuroticism.

Supplementary Table S1

Descriptive statistics and correlations among study variables (Study 1)

	<i>M</i>	<i>SD</i>	<i>N</i>	<i>Scale</i>	α	1	2	3	4	5	6	7	8	9	10	11	12	13
<i>Demographic background</i>																		
1. Sex (Male=1)	0.65	0.48	897	0-1	-													
2. Age	28.40	2.63	862	21-39	-	.19 **												
3. Non-native English	0.39	0.49	880	0-1	-	-.01	.26 **											
<i>Cognitive Intelligence</i>																		
4. Verbal GMAT	42.08	6.09	703	32-60	-	.03	-.09 *	-.27 **										
5. Quantitative GMAT	47.75	4.11	707	21-60	-	.04	.05	.28 **	.08 *									
6. Total GMAT	715.45	33.03	874	610-790	-	.07	-.01	-.03	.49 **	.39 **								
<i>Dispositional trait measures</i>																		
7. Positive affect	5.34	0.88	892	1-7	.85	-.12 **	.01	.03	-.02	-.01	-.04							
8. Negative affect	3.17	1.14	892	1-7	.88	-.06	-.11 **	-.05	.02	-.03	.02	-.31 **						
9. Openness	5.40	0.85	891	1-7	.76	.01	.12 **	.04	.04	-.02	.04	.36 **	-.17 **					
10. Conscientiousness	5.52	0.91	892	1-7	.79	-.06	.06	-.02	-.05	-.05	-.04	.21 **	-.28 **	.13 **				
11. Extroversion	4.97	1.14	893	1-7	.83	-.08 *	-.08 *	-.05	.00	-.04	-.09 *	.43 **	-.24 **	.30 **	.17 **			
12. Agreeableness	5.45	0.85	890	1-7	.70	-.16 **	.13 **	.09 *	-.13 **	-.04	-.13 **	.32 **	-.29 **	.19 **	.23 **	.14 **		
13. Neuroticism	3.21	1.08	892	1-7	.77	-.17 **	-.08 *	.00	-.01	.01	-.02	-.35 **	.79 **	-.19 **	-.27 **	-.24 **	-.30 **	
<i>Self-rated Emotional Intelligence</i>																		
14. Self-emotions appraisal	6.06	0.74	920	1-7	.84	-.02	.03	.04	-.06	.03	-.05	.17 **	-.25 **	.11 **	.14 **	.20 **	.18 **	-.24 **
15. Other-emotions appraisal	5.74	0.89	920	1-7	.86	-.11 **	-.01	.05	-.10 **	-.02	-.06	.17 **	.00	.15 **	.05	.12 **	.28 **	-.05
16. Use of emotion	6.18	0.69	920	1-7	.76	.05	-.05	-.02	.00	.00	.00	.19 **	-.20 **	.11 **	.37 **	.15 **	.09 **	-.23 **
17. Regulation of emotion	5.52	1.12	920	1-7	.90	.05	.08 *	.00	-.01	-.04	-.02	.22 **	-.41 **	.13 **	.12 **	.08 *	.33 **	-.49 **
18. Total EI	5.88	0.61	920	1-7	.87	-.01	.03	.02	-.06	-.01	-.05	.27 **	-.32 **	.18 **	.22 **	.18 **	.33 **	-.38 **
<i>Observer-rated Emotional Intelligence</i>																		
19. Self-emotions appraisal	5.84	1.45	931	a	.92	-.01	-.04	-.14 **	.11 **	-.13 **	-.01	.13 **	-.10 **	.00	.04	.09 **	.06	-.12 **
20. Other-emotions appraisal	5.44	2.43	931	a	.93	-.05	.01	-.08 *	-.02	-.12 **	-.07 *	.12 **	-.06	-.04	-.01	-.02	.14 **	-.06
21. Use of emotion	6.03	1.91	931	a	.84	.10 **	-.05	-.13 **	.15 **	.04	.15 **	.05	-.07 *	-.06	.15 **	.01	-.02	-.09 **
22. Regulation of emotion	5.69	2.22	931	a	.93	.16 **	.06	.04	.04	.02	.02	.06	-.15 **	-.04	-.04	-.12 **	.17 **	-.19 **
23. Total EI	5.75	1.52	931	a	.93	.07 *	.00	-.09 **	.08 *	-.06	.02	.12 **	-.12 **	-.05	.04	-.03	.13 **	-.15 **
<i>Performance ratings from Teammates</i>																		
24. Teammate	3.94	0.54	931	b	-	.10 **	-.09 *	-.21 **	.17 **	-.05	.10 **	.01	-.06	-.09 **	.10 **	-.06	-.02	-.09 **
25. Leadership	3.88	0.56	931	b	-	.17 **	-.10 **	-.29 **	.22 **	-.11 **	.08 *	.08 *	-.10 **	-.05	.10 **	.10 **	-.08 *	-.13 **
26. Individual effectiveness	3.38	0.57	931	b	-	.11 **	-.09 **	-.19 **	.16 **	.00	.16 **	-.03	-.02	-.09 **	.10 **	-.08 *	-.07 *	-.04

Supplementary Table S1, Continued.

Descriptive statistics and correlations among study variables (Study 1)

	14	15	16	17	18	19	20	21	22	23	24	25
<i>Self-rated Emotional Intelligence</i>												
14. Self-emotions appraisal												
15. Other-emotions appraisal	.42 **											
16. Use of emotion	.33 **	.20 **										
17. Regulation of emotion	.36 **	.32 **	.34 **									
18. Total EI	.71 **	.69 **	.61 **	.78 **								
<i>Observer-rated Emotional Intelligence</i>												
19. Self-emotions appraisal	.08 *	.08 *	.05	.09 **	.11 **							
20. Other-emotions appraisal	.06	.14 **	.00	.12 **	.12 **	.65 **						
21. Use of emotion	.02	-.06	.24 **	.03	.06 *	.37 **	.29 **					
22. Regulation of emotion	.05	.06	-.02	.25 **	.14 **	.48 **	.59 **	.19 **				
23. Total EI	.06 *	.08 *	.08 *	.17 **	.15 **	.79 **	.86 **	.59 **	.78 **			
<i>Performance ratings from Teammates</i>												
24. Teammate	.02	.00	.16 **	.04	.07 *	.51 **	.51 **	.74 **	.34 **	.68 **		
25. Leadership	.01	-.02	.16 **	.01	.05	.55 **	.48 **	.61 **	.29 **	.62 **	.75 **	
26. Individual effectiveness	.00	-.03	.19 **	-.01	.04	.39 **	.36 **	.78 **	.24 **	.57 **	.88 **	.65 **

Supplementary Table S2

Baseline Regression Models before Examining the Predictive Validity of Observer-reported Emotional Intelligence (Studies 1, 2, and 3)

	Study 1						Study 2						Study 3							
	Effectiveness as a		Effectiveness as a		Effectiveness as an		OCB		Transformational		Job		Effectiveness as a				Effectiveness as an			
	Teammate		Leader		Individual				Leadership		Performance		Teammate				Individual			
Intercept	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Age	.10 **	.09 *	.17 **	.17 **	.12 **	.11 **	.03	.02	-.02	-.04	-.03	-.04	.03	-.01	.02	-.01	-.01	-.05	-.01	-.04
Male	-.06	-.05	-.04	-.04	-.07	-.06	-.11	-.10	-.07	-.03	-.02	.05	-.09	-.10 **	-.07	-.09 *	-.05	-.07	-.02	-.04
Non-native English speaker	-.18 **	-.18 **	-.26 **	-.26 **	-.16 **	-.16 **	-.05	-.01	-.06	-.06	-.18 *	-.17	-.05	-.03	-.06	-.04	.00	.01	.00	.01
Positive Affect	.06	.06	.11 **	.10 **	.04	.04	.15	.16	.23 *	.15	.28 **	.15	.02	.02	.04	.04	.02	.02	.07	.07
Negative Affect	.00	.00	-.02	-.02	-.01	-.01	-.10	-.11	-.09	-.10	-.05	-.08	-.01	-.06	.01	-.05	-.02	-.07	-.03	-.08
Extraversion	-.10 **	-.09 *	.07	.08 *	-.11 **	-.10 **	-.11	-.15	-.03	.04	.01	.19	.06	-.02	.06	-.03	.09	.02	.08	.01
Conscientiousness	.10 **	.10 **	.09 *	.09 *	.13 **	.14 **	-.19 *	-.17	-.21 *	-.06	-.11	.01	.18 *	.09 *	.20 **	.10 *	.18 *	.10 *	.19 **	.11 *
Openness	-.09 *	-.10 **	-.10 **	-.10 **	-.08 *	-.09 *	.02	-.11	-.05	-.20	-.15	-.26 *	-.05	.03	.00	.05	.03	.10 *	.06	.10 *
Agreeable	-.01	.00	-.08 *	-.07	-.05	-.03	.13	.08	.10	.09	.00	-.08	-.06	-.01	-.08	-.01	-.11	-.06	-.12	-.07
Neuroticism	-.08	-.08	-.06	-.05	-.04	-.03	.14	.12	.14	.11	.12	.06	.07	.05	.08	.06	.07	.06	.10	.08
GMAT score		.09 *		.08 *		.13 **		-.18		-.02		.17								
SAT score														.03	.00				.14 *	.12 **
Observer-rated Liking/Regard													.80 **		.80 **		.68 **		.67 **	
DF	796	795	796	795	796	795	124	85	124	85	124	85	272	271	254	253	272	271	254	253
R-Squared	.08	.09	.15	.16	.08	.10	.09	.10	.10	.07	.10	.15	.05	.67	.06	.67	.06	.51	.10	.53
Adj. R-Squared	.07	.08	.14	.14	.07	.09	.02	-.01	.03	-.05	.02	.03	.02	.65	.02	.66	.02	.49	.06	.51

*p < .05; **p<.01; all values two-tailed.

Notes. Multiple models are included to illustrate results with and without proxy for cognitive intelligence (GMAT or SAT), and for Study 3 with and without observer-rated liking/regard.

Supplementary Table S3
Descriptive statistics and correlations among study variables (Study 2)

	M	SD	N	Scale	α	1	2	3	4	5	6	7	8	9	10	11
<i>Demographic background</i>																
1. Male	0.71	0.45	135	0-1	-											
2. Age	29.90	5.44	135	23-56	-	.14										
3. Non-native English Speaker	0.13	0.34	135	0-1	-	.11	.21 *									
<i>Cognitive Intelligence</i>																
4. Total GMAT	614.18	66.18	97	410-760	-	.10	.17	.07								
<i>Dispositional trait measures</i>																
5. Positive affect	3.81	0.49	135	1-7	.83	-.13	.02	.10	-.06							
6. Negative affect	1.76	0.45	135	1-7	.76	.09	.08	-.08	.07	-.23 **						
7. Openness	5.49	0.88	135	1-7	.52	-.06	-.09	-.03	-.09	.37 **	-.17 *					
8. Conscientiousness	6.00	0.94	135	1-7	.66	-.17	-.13	.00	-.05	.16	-.21 *	-.02				
9. Extroversion	4.90	1.24	135	1-7	.52	-.14	.08	-.02	-.15	.33 **	-.09	.38 **	-.13			
10. Agreeableness	5.03	0.99	135	1-7	.26	-.20 *	-.03	.11	.02	.25 **	-.24 **	.13	.17	.08		
11. Neuroticism	3.20	1.30	135	1-7	.25	-.07	.02	.01	-.03	-.22 **	.25 **	-.12	-.03	.09	-.15	
<i>Self-rated Emotional Intelligence</i>																
12. Perceiving emotion	7.00	1.04	135	1-9	.74	-.01	-.05	.02	-.12	.37 **	-.26 **	.20 *	.09	.31 **	.15	-.12
13. Use of emotion	5.20	1.58	135	1-9	.87	-.27 **	.05	.00	.08	.14	.03	.19 *	.05	.24 **	.30 **	.13
14. Understanding emotion	5.65	1.66	135	1-9	.87	-.24 **	-.04	-.20 *	.13	.12	-.19 *	.25 **	-.05	.25 **	.18 *	-.02
15. Managing	6.58	0.93	135	1-9	.67	.07	-.13	-.08	-.17	.47 **	-.45 **	.35 **	.11	.17 *	.27 **	-.26 **
16. Total EI	6.30	0.80	135	1-9	.85	-.11	-.08	-.11	-.06	.42 **	-.39 **	.37 **	.06	.34 **	.32 **	-.14
<i>Observer-rated Emotional Intelligence</i>																
17. Perceiving emotion	7.26	0.56	135	1-9	.78	-.19 *	-.19 *	-.08	-.07	.10	-.07	-.06	-.09	-.04	.29 **	-.06
18. Use of emotion	4.75	0.64	135	1-9	.57	-.24 **	.09	-.04	-.07	-.04	.08	.05	.00	.20 *	.10	.01
19. Understanding emotion	7.02	0.68	135	1-9	.82	-.33 **	-.07	-.08	-.04	.02	-.05	.01	-.06	.05	.17	.01
20. Managing	7.01	0.67	135	1-9	.74	-.09	-.24 **	.00	-.04	.16	-.12	-.04	.05	-.16	.31 **	-.10
21. Total EI	6.80	0.46	135	1-9	.86	-.25 **	-.21 *	-.05	-.06	.13	-.10	-.03	.00	-.06	.35 **	-.08
<i>Ability-tested Emotional Intelligence</i>																
22. Perceiving emotion	97.75	14.31	135	62-129	-	.05	.03	-.02	.04	-.06	-.18 *	-.06	.13	.05	.06	-.01
23. Use of emotion	97.35	13.79	135	63-132	-	.00	.02	.10	.02	-.05	-.24 **	-.01	.05	.06	.16	.02
24. Understanding emotion	101.75	8.78	135	78-120	-	-.19 *	.00	-.21 *	.24 *	-.16	-.12	-.11	-.08	-.04	-.01	-.05
25. Managing	96.80	7.94	135	75-114	-	-.03	-.08	-.16	-.14	.23 **	-.25 **	.25 **	.18 *	.17	.15	-.18 *
26. Total EI	98.53	11.76	135	69-123	-	-.04	-.01	-.09	.02	-.04	-.25 **	.01	.10	.10	.13	-.05
<i>Performance ratings from Supervisors</i>																
27. Organizational citizenship behavior	5.52	0.68	135	1-7	-	-.13	.01	-.02	-.17	.12	-.08	.06	-.10	.02	.15	.06
28. Transformational leadership	5.50	0.41	135	1-7	-	-.11	-.01	-.03	-.04	.19 *	-.08	.05	-.12	.09	.12	.06
29. Performance	5.98	0.56	135	1-7	-	-.06	-.04	-.15	.14	.18 *	-.02	-.04	-.05	.08	.01	.06

*p < .05; **p < .01; all values two-tailed.

Supplementary Table S3, Continued.
 Descriptive statistics and correlations among study variables (Study 2)

	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
<i>Self-rated Emotional Intelligence</i>																	
12. Perceiving emotion																	
13. Use of emotion	.22 *																
14. Understanding emotion	.37 **	.30 **															
15. Managing	.45 **	.01	.25 **														
16. Total EI	.73 **	.38 **	.75 **	.74 **													
<i>Observer-rated Emotional Intelligence</i>																	
17. Perceiving emotion	.17 *	.14	.05	.09	.15												
18. Use of emotion	.07	.19 *	.14	-.14	.06	-.01											
19. Understanding emotion	.09	.20 *	.27 **	.06	.21 *	.50 **	.21 *										
20. Managing	.04	.07	-.03	.28 **	.15	.64 **	-.27 **	.33 **									
21. Total EI	.13	.19 *	.11	.19 *	.22 *	.83 **	.08	.70 **	.85 **								
<i>Ability-tested Emotional Intelligence</i>																	
22. Perceiving emotion	.06	.00	.14	.03	.09	-.07	.08	.07	-.11	-.04							
23. Use of emotion	.11	-.01	.00	.12	.08	-.02	.00	.04	-.03	.00	.61 **						
24. Understanding emotion	.02	-.10	.15	-.02	.04	.02	-.01	.16	-.01	.05	.28 **	.26 **					
25. Managing	.27 **	.04	.13	.30 **	.28 **	.10	.17	.16	.04	.14	.25 **	.34 **	.27 **				
26. Total EI	.15	.00	.15	.13	.16	-.01	.08	.14	-.05	.04	.84 **	.80 **	.55 **	.57 **			
<i>Performance ratings from Supervisors</i>																	
27. Organizational citizenship behavior	.07	.10	.21 *	.11	.20 *	.22 *	.13	.17 *	.28 **	.32 **	-.17 *	-.12	.01	-.05	-.12		
28. Transformational leadership	.04	.08	.17	.06	.15	.31 **	.05	.16	.34 **	.37 **	-.15	-.16	.04	-.13	-.15	.72 **	
29. Performance	.01	.01	.13	.00	.07	.14	.08	.10	.18 *	.20 *	-.11	-.21 *	.07	-.07	-.11	.52 **	.72 **

Supplementary Table S4

Results of a Confirmatory Factor Analysis examining Self-rated, Observer-rated, and Ability-tested Emotional Intelligence (Study 2)

	Method Loadings		Trait Loadings		Residual Variance		R-Squared
	Estimate	Std. Err	Estimate	Std. Err	Estimate	Std. Err	
<i>Self-rated Emotional Intelligence</i>							
Perceiving emotion	0.69 **	(0.14)	0.42 **	(0.15)	0.35	(0.15)	65%
Use of emotion	0.16 **	(0.12)	0.40 **	(0.11)	0.81	(0.12)	19%
Understanding emotion	0.44 **	(0.11)	0.52 **	(0.12)	0.52	(0.11)	48%
Managing	0.45 **	(0.12)	0.85 **	(0.27)	0.05	(0.45)	95%
<i>Observer-rated Emotional Intelligence</i>							
Perceiving emotion	0.75 **	(0.12)	0.51 **	(0.15)	0.16	(0.16)	84%
Use of emotion	-0.14	(0.11)	0.49 **	(0.13)	0.74	(0.14)	26%
Understanding emotion	0.40 **	(0.10)	0.61 **	(0.12)	0.44	(0.12)	56%
Managing	0.68 **	(0.11)	0.33 *	(0.14)	0.42	(0.11)	58%
<i>MSCEIT</i>							
Perceiving emotion	0.73 **	(0.10)	-0.14	(0.11)	0.44	(0.11)	56%
Use of emotion	0.81 **	(0.10)	-0.16	(0.10)	0.33	(0.12)	67%
Understanding emotion	0.35 **	(0.10)	0.08	(0.11)	0.86	(0.11)	14%
Managing	0.41 **	(0.09)	0.25 *	(0.12)	0.76	(0.11)	24%

Notes:

N=135. Model fit statistics: *Chi-square*(33)=37.95, $p=.25$, baseline *chi-square*(66)=345.03, $p<.01$. CFI=.98, TFI=.96, log likelihood=-2053.98, RMSEA = .03, 95% CI .00-.08, $p=.69$.

Supplementary Table S5a

Partial correlations between Emotional Intelligence and Performance Ratings, Controlling for Demographic Variables, Personality Traits, and Cognitive Intelligence (Studies 2 and 3)

	Study 2			Study 3			
	Performance Ratings from Supervisors			Performance Ratings from Teammates		Performance Ratings from Teammates Controlling for Liking	
	Organizational Citizenship	Transformational Leadership	Job Performance	Individual Effectiveness	Individual Teammate	Individual Effectiveness	Teammate Effectiveness
a. Without controlling for Cognitive Intelligence							
	N=135			N=283			
Self-Rated EI							
Perceiving emotion	.04	-.04	-.06	-.01	.03	-.08	-.06
Use of emotion	-.05	-.11	-.08	-.09	-.05	-.08	-.03
Understanding emotion	.16	.11	.10	.03	.03	-.02	-.02
Managing emotion total	.06	-.06	-.13	.03	.05	-.04	-.04
Total EI	.13	.01	-.03	-.02	-.02	-.04	-.04
CI for total EI	(-.08, .35)	(-.21, .22)	(-.25, .19)	(-.18, .13)	(-.17, .13)	(-.15, .07)	(-.13, .05)
With double-barreled questions:							
Use of emotion	.03	.01	-.02	-.10	-.03	-.13 ** (.01)	-.06
Total EI	.16	.05	-.02	.00	.04	-.09	-.07
CI for total EI	(-.06, .38)	(-.17, .26)	(-.24, .20)	(-.13, .13)	(-.09, .17)	(-.19, .00)	(-.15, .01)
Observer-Rated EI							
Perceiving emotion	.16	.27 ** (.06)	.09	.49 ** (.22)	.65 ** (.40)	.00	.17 ** (.01)
Use of emotion	.22 * (.04)	.21 * (.04)	.12	.22 ** (.04)	.30 ** (.08)	-.07	-.03
Understanding emotion	.11	.11	.07	.50 ** (.21)	.57 ** (.27)	.17 ** (.02)	.18 ** (.02)
Managing emotion total	.26 ** (.05)	.34 ** (.09)	.17	.45 ** (.20)	.65 ** (.42)	-.05	.19 ** (.02)
Total EI	.26 ** (.05)	.34 ** (.09)	.16	.52 ** (.26)	.70 ** (.47)	.02	.24 ** (.02)
CI for total EI	(.07, .45)	(.16, .53)	(-.03, .36)	(.42, .62)	(.61, .79)	(-.11, .16)	(.13, .34)
With double-barreled questions:							
Use of emotion	.14	.05	.11	-.22 ** (.04)	-.17 ** (.02)	-.15 ** (.02)	-.09 * (.01)
Total EI	.29 ** (.07)	.37 ** (.10)	.19 * (.03)	.49 ** (.22)	.68 ** (.43)	-.02	.20 ** (.02)
CI for total EI	(.10, .48)	(.18, .55)	(.00, .39)	(.38, .60)	(.59, .77)	(-.14, .11)	(.10, .30)
Ability-tested EI							
Perceiving emotion	-.16	-.14	-.11	-	-	-	-
Use of emotion	-.15	-.19 * (.03)	-.21 * (.04)	-	-	-	-
Understanding emotion	-.02	.03	.05	-.02	-.02	.01	.01
Managing emotion total	-.08	-.19 * (.03)	-.12	.09	.02	.13 ** (.01)	.07
Total EI	-.15	-.17	-.13	-	-	-	-
CI for total EI	(-.33, .04)	(-.35, .01)	(-.31, .05)	-	-	-	-

Supplementary Table S5b

Partial correlations between Emotional Intelligence and Performance Ratings, Controlling for Demographic Variables, Personality Traits, and Cognitive Intelligence (Studies 2 and 3)

	Study 2			Study 3			
	Performance Ratings from Supervisors			Performance Ratings from Teammates		Performance Ratings from Teammates Controlling for Liking	
	Organizational Citizenship	Transformational Leadership	Job Performance	Individual Effectiveness	Individual Teammate	Individual Effectiveness	Teammate Effectiveness
	N=97			N=266			
Self-Rated EI							
Perceiving emotion	.10	.12	.02	.02	.04	-.07	-.06
Use of emotion	-.12	-.01	-.09	-.05	-.03	-.05	-.03
Understanding emotion	.17	.20	.16	.04	.06	-.01	-.01
Managing emotion total	.11	-.02	-.05	.06	.05	-.03	-.05
Total EI	.17	.15	.07	.03	.02	.00	-.02
CI for total EI	(-.08, .42)	(-.09, .40)	(-.12, .29)	(-.12, .19)	(-.13, .18)	(-.12, .11)	(-.11, .07)
With double-barreled questions:							
Use of emotion	.00	.10	.03	-.07	-.02	-.13 *	(.01) -.08
Total EI	.19	.20	.10	.04	.06	-.08	-.08
CI for total EI	(-.06, .44)	(-.05, .44)	(-.12, .33)	(-.10, .17)	(-.08, .20)	(-.18, .02)	(-.16, .00)
Observer-Rated EI							
Perceiving emotion	.16	.27 ** (.08)	.08	.52 ** (.25)	.68 ** (.43)	.06	.20 ** (.02)
Use of emotion	.17	.18	.10	.24 ** (.05)	.29 ** (.08)	-.05	-.04
Understanding emotion	.08	.12	.08	.50 ** (.21)	.57 ** (.29)	.17 ** (.02)	.18 ** (.02)
Managing emotion total	.19	.26 * (.06)	.09	.48 ** (.22)	.66 ** (.44)	.00	.21 ** (.02)
Total EI	.18	.26 * (.07)	.09	.54 ** (.28)	.71 ** (.49)	.08	.26 ** (.03)
CI for total EI	(-.03, .39)	(.06, .46)	(-.09, .28)	(.44, .64)	(.62, .79)	(-.06, .21)	(.15, .36)
With double-barreled questions:							
Use of emotion	.01	-.07	.05	-.20 ** (.03)	-.18 * (.02)	-.13 ** (.01)	-.10 * (.01)
Total EI	.19	.27 ** (.08)	.10	.51 ** (.25)	.69 ** (.45)	.04	.22 ** (.02)
CI for total EI	(-.02, .41)	(.07, .48)	(-.09, .29)	(.41, .62)	(.60, .78)	(-.10, .17)	(.12, .32)
Ability-tested EI							
Perceiving emotion	-.15	-.14	-.03	-	-	-	-
Use of emotion	.00	-.12	-.11	-	-	-	-
Understanding emotion	.01	-.02	-.06	-.04	.00	.00	.04
Managing emotion total	-.07	-.23	-.19	.06	.04	.11 * (.01)	.10 * (.01)
Total EI	-.10	-.19	-.11	-	-	-	-
CI for total EI	(-.32, .13)	(-.40, .03)	(-.31, .08)	-	-	-	-

Notes:

For significant coefficients, the incremental R-squared appears in parentheses

*p < .05; **p < .01.

Supplementary Table S6

Descriptive statistics and correlations among study variables (Study 3)

	<i>M</i>	<i>SD</i>	<i>N</i>	<i>Scale</i>	α	1	2	3	4	5	6	7	8	9	10	11	12
<i>Demographic background</i>																	
1. Male	0.58	0.49	287	0-1	-												
2. Age	19.92	1.02	287	18-24	-	.10											
3. Native English Speaker	0.78	0.42	285	0-1	-	.20 **	.01										
<i>Cognitive Intelligence</i>																	
4. Total SAT	2175.6	131.5	269	1740-2400	-	-.01	-.04	.01									
<i>Dispositional trait measures</i>																	
5. Positive affect	3.55	0.64	292	1-5	.87	.06	-.07	.18 **	-.08								
6. Negative affect	2.04	0.62	292	1-5	.84	-.02	-.01	-.07	.07	-.10							
7. Openness	5.10	1.08	292	1-7	.48	-.07	-.02	.10	-.06	.32 **	-.21 **						
8. Conscientiousness	5.57	1.17	292	1-7	.58	-.03	-.15 *	.05	-.06	.45 **	-.24 **	.10					
9. Extroversion	4.73	1.32	292	1-7	.74	-.08	.00	.12 *	.08	.33 **	-.07	.26 **	-.01				
10. Agreeableness	4.91	1.02	292	1-7	.25	-.15 **	-.05	-.16 **	.00	.12 *	-.23 **	.11	.17 **	-.05			
11. Neuroticism	2.87	1.21	292	1-7	.64	-.23 **	.04	-.04	.03	-.28 **	.52 **	-.15 *	-.22 **	.03	-.24 **		
<i>Ability-rated Emotional Intelligence</i>																	
12. STEM	4.41	0.27	289	3.3-4.9	.85	-.16 **	-.11	.08	.05	.14 *	-.14 *	.27 **	.08	.09	.12 *	-.11	
13. STEU	0.69	0.12	292	.19-.90	.69	-.07	-.10	.07	.09	.03	-.13 *	.17 **	.05	.02	.15 *	-.07	.59 **
<i>Self-rated Emotional Intelligence</i>																	
14. Perceiving Emotions	6.95	1.12	299	1-9	.72	.02	-.03	.14 *	-.03	.18 **	-.11	.23 **	.17 **	.22 **	.11	-.05	-.03
15. Understanding Emotions	5.03	1.57	299	1-9	.94	-.21 **	-.03	.00	-.07	-.10	.04	.07	-.17 **	.10	.23 **	.27 **	.03
16. Use of Emotions	6.21	1.70	299	1-9	.75	-.09	-.04	.11	.01	.09	-.06	.28 **	.06	.08	.12 *	.04	.10
17. Managing Emotions	6.75	1.03	299	1-9	.75	.09	-.06	.11	-.12	.29 **	-.25 **	.26 **	.24 **	.19 **	.11	-.32 **	.13 *
18. Total EI	6.41	0.84	299	1-9	.82	-.05	-.07	.14 *	-.09	.21 **	-.17 **	.34 **	.15 *	.23 **	.21 **	-.08	.11
<i>Observer-rated Emotional Intelligence</i>																	
19. Perceiving Emotions	5.81	0.64	299	a	.83	-.12 *	.08	.09	.04	.02	-.01	-.02	.03	.11	.06	.04	-.04
20. Understanding Emotions	4.87	0.63	299	a	.89	-.26 **	-.17 **	.07	-.07	.01	.10	.03	-.11	.13 *	.02	.21 **	.08
21. Use of Emotions	5.79	0.64	299	a	.67	-.23 **	.01	.11	.09	.05	-.04	.07	.00	.20 **	.08	.07	.11
22. Managing Emotions	5.94	0.68	299	a	.87	.08	.14 *	.06	.00	.01	-.01	-.04	-.01	.04	.02	-.10	-.04
23. Total EI	5.71	0.50	299	a	.90	-.10	.07	.10	.02	.02	.00	-.01	-.02	.13 *	.05	.01	.01
<i>Performance ratings from Teammates</i>																	
24. Liking	4.99	0.78	299	b	.92	.03	.05	.04	.04	.03	.08	-.09	.07	.07	-.07	.05	-.06
25. Effectiveness as an individual	5.47	0.97	299	b	-	-.07	-.03	.03	.13 *	.10	-.01	.06	.18 **	.11	-.06	.05	.02
26. Effectiveness as a teammate	5.45	0.92	299	b	-	-.09	.00	.05	.02	.08	.01	-.03	.17 **	.05	-.02	.06	.01

*p < .05; **p < .01; all values two-tailed.

a Values listed in descriptive statistics are on the original scale of 1-9. For analysis, peer-rated EI values generated by SOREMO mean-centered by group.

b Values listed in descriptive statistics are on the original scale of 1-7. For analysis, peer-rated values generated by SOREMO mean-centered by group.

Supplementary Table S6, Continued.

Descriptive statistics and correlations among study variables (Study 3)

	13	14	15	16	17	18	19	20	21	22	23	25	26
<i>Self-rated Emotional Intelligence</i>													
14. Perceiving Emotions	.03												
15. Understanding Emotions	.06	.04											
16. Use of Emotions	.08	.37 **	.23 **										
17. Managing Emotions	.07	.52 **	-.08	.27 **									
18. Total EI	.10	.71 **	.36 **	.74 **	.75 **								
<i>Observer-rated Emotional Intelligence</i>													
19. Perceiving Emotions	-.04	.14 *	.13 *	.11 *	.12 *	.19 **							
20. Understanding Emotions	-.03	-.03	.27 **	.06	-.05	.07	.12 *						
21. Use of Emotions	.12 *	.02	.17 **	.11	.06	.13 *	.68 **	.21 **					
22. Managing Emotions	-.06	.07	-.04	-.03	.15 **	.07	.74 **	-.09	.52 **				
23. Total EI	-.02	.07	.11	.06	.12 *	.14 *	.89 **	.24 **	.79 **	.89 **			
<i>Performance ratings from Teammates</i>													
24. Liking	-.06	.08	.05	.04	.08	.09	.69 **	-.04	.49 **	.69 **	.70 **		
25. Effectiveness as an individual	.11	.04	-.06	.06	.08	.06	.49 **	-.13 *	.48 **	.42 **	.47 **	.68 **	
26. Effectiveness as a teammate	.04	.04	.00	.06	.06	.07	.66 **	-.08	.55 **	.62 **	.66 **	.80 **	.85 **

* $p < .05$; ** $p < .01$; all values two-tailed.

a Values listed in descriptive statistics are on the original scale of 1-9. For analysis, peer-rated EI values generated by SOREMO mean-centered by group.

b Values listed in descriptive statistics are on the original scale of 1-7. For analysis, peer-rated values generated by SOREMO mean-centered by group.