Does emotional intelligence as ability predict transformational leadership? A multilevel approach ☆,☆☆

Hallvard Føllesdal [a,b,*], Knut Hagtvet [b]

[a] Department of Leadership and Organizational Behaviour, BI Norwegian Business School, Norway
[b] Department of Psychology, University of Oslo, Norway

ARTICLE INFO

ABSTRACT

This study assesses to what extent emotional intelligence as ability (EI) can predict transformational leadership. Norwegian executives (N = 104) completed measures of personality (NEO PI-R) and EI (MSCEIT), and were rated on transformational leadership (MLQ 5X) by subordinates (N = 459). This study improves upon previous studies in three ways: First, because the validity and reliability of the scores from MSCEIT has been questioned, an alternative set of scales from MSCEIT were included, which provide reliable and interpretable scores (Føllesdal & Hagtvet, 2009). Second, in addition to personality, general mental ability (GMA) was controlled for by utilizing Monte Carlo studies. Third, a multilevel approach was used to analyze the scores, due to their hierarchical structure. Neither the four branch scores, nor the Total EI score from MSCEIT predicted transformational leadership. A suppression effect was found, however, among two subscales from Perceiving Emotions. The validity of scores from MSCEIT is questioned.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Emotional intelligence (EI) is assumed to be an important characteristic in leadership, particularly in transformational leadership (Ashkanasy & Tse, 2000; George, 2000). There exist many different conceptualizations of EI, e.g., as a set of abilities (Mayer & Salovey, 1997), a set of personality traits (Bar-On, 1997), or a mixture of effective leader behaviors and personality traits (Boyatzis, Goleman, & Rhee, 2000). Many researchers, however, argue that EI should be conceptualized and measured as a set of abilities, distinct from personality traits and leader behavior. A recent meta-analysis, however, found that EI measured as ability was unrelated to subordinates' ratings of EI (Harms & Credé, 2010), suggesting that EI as ability might not be important in transformational leadership. There are three limitations with previous research that will be addressed in the present study, in order to assess more thoroughly the relationship between EI as ability and transformational leadership.

First, EI as ability is usually measured by the Mayer, Salovey, Caruso Emotional Intelligence Test (MSCEIT; Mayer, Salovey, & Caruso, 2002), which is the only performance test measuring the entire four-branch ability model of EI. Many of the scales from this test, however, provide scores with low reliability, that are difficult to interpret, and with questionable validity (Fiori & Antonakis, 2011; Føllesdal & Hagtvet, 2009; Keele & Bell, 2008; Palmer, Gignac, Manocha, & Stough, 2005; Rode et al., 2008; Rossen, Kranzler, & Algina, 2008). The use of these scales may therefore lead to inconsistent and unexpected findings, e.g., that EI as ability is unrelated to transformational leadership. Føllesdal and Hagtvet (2009), however, identified a set of alternative scales in the MSCEIT that provide reliable and interpretable scores. The present study will use these scales, in addition to the standard scales from MSCEIT, in order to assess the relationship with transformational leadership.

Second, studies assessing the relationship between EI as ability and transformational leadership have seldom controlled for both personality traits and general mental ability (GMA). This is necessary in order to demonstrate that EI is a unique characteristic that is important in leadership (see e.g., Antonakis, Ashkanasy, & Dasborough, 2009; Walter, Cole, & Humphrey, 2011). The present study will therefore control for the personality factors in the Five Factor Model of personality (FFM) and GMA when assessing the relationship between EI as ability and transformational leadership. GMA will be controlled for by the use of Monte Carlo studies.

Third, studies assessing the relationship between EI as ability and transformational leadership seldom take into account the multilevel nature of scores from measures of transformational leadership. That is, transformational leadership is often measured by subordinates' ratings of their leader, which introduces dependencies in the scores. Ignoring such dependencies may lead to biased parameter estimates and wrong conclusions about how the scores relate to other variables (Bliese & Hanges, 2004; Hox, 2002), such as leaders' EI. The present study will assess the relationship between EI as ability and transformational leadership within a multilevel framework, which takes such dependencies into account.

In the following, we will first describe transformational leadership and the theoretical relationship with EI, and some empirical findings regarding this relationship. Second, we will look at some issues that may question the validity of the scores from MSCEIT, and explain why it may be appropriate to assess the predictive validity of an alternative set of scales from MSCEIT. Third, we will explain the need for a multilevel framework when assessing the relationship between the scores on transformational leadership and EI as ability. Finally, we will report and discuss results from a study assessing the relationship between scores from MSCEIT and transformational leadership, in a sample of 104 Norwegian executives rated on transformational leadership by subordinates, controlling for the FFM and GMA.

2. Transformational leadership

Transformational leadership is a type of leader behavior that consists of four components (Bass, 1997; Bass & Riggio, 2006): *Idealized Influence* describes leaders who behave in ways that allow them to serve as role models for their subordinates and who are admired, respected, and trusted. *Inspirational Motivation* describes leaders who behave in ways that motivate and inspire subordinates, e.g., by providing meaning and challenge to their work, and establishing inspiring visions. *Intellectual Stimulation* describes leaders who stimulate their followers to be innovative and creative by questioning assumptions, reframing problems, and approaching situations in new ways. Finally, *Individualized Consideration* describes leaders who create new learning opportunities along with a supportive climate, who recognize individual differences in terms of needs and desires (Bass, 1997; Bass & Riggio, 2006).

Researchers have tried to identify individual differences that may explain variance in transformational leadership. A meta-analysis reported that the five broad personality factors in the FFM explain about 9% of the variance in transformational leadership (Bono & Judge, 2004). When it comes to GMA, a recent meta-analysis reported a correlation of .16 (corrected for unreliability) between transformational leadership and GMA, suggesting that GMA only explain 2–3% of the variance in transformational leadership (DeRue, Nahrgang, Wellman, & Humphrey, 2011). This suggests that neither the FFM nor GMA are important predictors of transformational leadership. Several researchers, however, assume that EI might be an important predictor of transformational leadership.

3. Emotional intelligence

There exist many different conceptualizations of EI, but the four-branch ability model (Mayer et al., 2002; Salovey & Sluyter, 1997) is often considered to be the most promising model of EI. In the four-branch model, EI is defined as four broad mental abilities assumed to meet traditional standards of an intelligence (Mayer, Caruso, & Salovey, 1999), i.e., ability to perceive, use, understand, and manage emotions (Mayer et al., 2002; Salovey & Sluyter, 1997).

3.1. Theoretical relationships between EI and transformational leadership

Several researchers have argued that EI as ability is important in effective leadership in general (Caruso, Mayer, & Salovey, 2002; Caruso & Salovey, 2004; George, 2000; Rosete & Ciarrochi, 2005) and in transformational leadership in particular (Ashkanasy & Tse, 2000; Bass, 2002; Bass & Riggio, 2006; Brown & Moshavi, 2005). All four branches in the four-branch model seem to be important in transformational leadership.

The first branch, Perceiving Emotions, encompasses ability to e.g., identify emotions in one's physical and psychological states and in other people, express emotions accurately, and express needs related to these emotions (Salovey & Mayer, 2002). These abilities may be important in transformational leadership because it might be difficult to inspire, motivate, stimulate and show individual consideration without the ability to accurately appraise and express emotions in self and others (Kupers & Weibler, 2006; Lindebaum & Cartwright, 2010).

The second branch, Using Emotions, encompasses e.g., the ability to generate emotions to facilitate judgment and memory, and use emotional states to facilitate problem-solving and creativity (Salovey & Mayer, 2002). Ability to use emotions is assumed to be linked to inspirational motivation (Lindebaum & Cartwright, 2010), and leaders who are able to generate and use emotion may be better able to generate positive emotions that facilitate creativity, which enables them to make more compelling visions (George, 2000).

The third branch, Understanding Emotions, encompasses e.g., ability to understand relationships among various emotions, and ability to perceive the causes and consequences of emotions (Salovey & Mayer, 2002). These abilities may be important in order to generate and maintain enthusiasm (George, 2000), which is an important aspect of Inspirational Motivation. And these abilities

may be important in order to anticipate how followers will react to different circumstances, events and changes (George, 2000), which might be important in Individualized Consideration.

The fourth branch, Managing Emotions, encompasses e.g., ability to manage own and others emotions. These abilities might be important in order to make followers aware of problems, yet make them confident about resolving the problems, and feel optimistic about the efficacy of their own contributions (George, 2000). Ashkanasy and Tse (2000) has also argued that managing emotions is important in transformational leadership in several ways.

One may thus argue that all four branches in the four-branch model of EI may be important in transformational leadership. It has been stated that "transformational leadership is intrinsically associated with emotional intelligence" (Ashkanasy & Tse, 2000, p. 232), and "the components of transformational leadership clearly resemble the key components of emotional intelligence" (Ashkanasy, Hartel, & Daus, 2002, p. 325). Other researchers, however, have questioned the role of EI in leadership (Antonakis et al., 2009).

3.2. The measurement of EI

EI is a relatively new construct in psychology, and as mentioned, several models and measures of EI have been developed. Some models encompass a set of personality traits (Bar-On, 1997, 2000; Petrides, 2010), and other models a mixture of personality traits and effective leadership behaviors (Boyatzis et al., 2000). Many researchers, however, has argued that EI should be measured with performance tests, and presently the MSCEIT is the only performance test that measures the entire four-branch model of EI, and thus has provided the most important empirical evidence for this model (for a review, see Mayer, Roberts, & Barsade, 2008).

3.3. Empirical studies of EI and transformational leadership

Despite the assumed positive relationship between EI and transformational leadership, a recent meta-analysis found that EI as ability, measured with the MSCEIT, is unrelated to subordinates ratings of transformational leadership (Harms & Credé, 2010). This may indicate that the constructs in the four branch ability model of EI are not important in transformational leadership, at least as perceived by their subordinates. An alternative explanation may be that the MSCEIT does not provide valid scores of the four-branch model of EI. The same meta-analysis, however, reported that when the four-branch model was measured using self-report, the scores predicted subordinates' ratings of transformational leadership. This finding may suggest that the MSCEIT might not be an optimal measure of the four-branch model of EI. This is in line with findings that the scores from MSCEIT, contrary to what is expected, do not predict job performance (Joseph & Newman, 2010; O'Boyle, Humphrey, Pollack, Hawver, & Story, 2011). We will therefore scrutinize the validity of the scores from MSCEIT, and suggest an alternative set of scales for some of the branches.

4. Validity issues related to the scores from MSCEIT

The MSCEIT measures each branch in the four-branch model by two tasks, consisting of several stimuli to be rated on a set of items. Each item response is weighted according to the correctness of that response. The test provides four branch scores and a Total EI score. Some studies support the validity of scores (for a discussion, see Mayer, Salovey, & Caruso, 2004a, 2004b; Mayer, Salovey, & Caruso, 2008; Mayer et al., 2002), while other studies have questioned their validity (e.g., Antonakis et al., 2009; Føllesdal & Hagtvet, 2009; Maul, 2012; Roberts et al., 2006).

There are several issues that may question the validity of the scores from MSCEIT. First, the scores have been found to contain large amounts of measurement error, and the reliability of the scores seems to be very low (Føllesdal & Hagtvet, 2009). Most studies, however, only report the reliability coefficients in the User's Manual, instead of estimating reliability coefficients in the actual sample of study. Thus, the reliability of scores may be much lower than reported, which may explain why reported relationships, e.g., with transformational leadership, is weak.

Second, only 122 of the 141 items in the MSCEIT are used in the scales (Mayer et al., 2002, p. 63), i.e., scores from 19 items are excluded before estimating reliability and calculating scale scores, based on results from initial exploratory factor analysis and weak item-scale correlations (G. Sitarenios, personal communication, December 11th, 2003). By using these scales, omitting scores from some items without understanding why they correlate negatively with other items measuring the same construct, one cannot be sure about what is measured. This makes interpretation of findings difficult.

Third, despite some studies that have reported good fit indices from Confirmatory Factor Analysis (CFA) of scores from the MSCEIT, they are not informative about the fit between the item scores and the theoretical model (e.g., Day & Carroll, 2004; Mayer, Salovey, Caruso, & Sitarenios, 2003). Instead of testing the correspondence between the 141 item scores and the underlying theoretical model, the item scores have been averaged into eight task scores, or parcels, which have been used in the analyses. The use of parceling may increase model fit for misspecified models and reduce our ability to identify the model as not fitting the data (Bandalos & Finney, 2001; Kim & Hagtvet, 2003). The technique of parceling requires that the scores are unidimensional (Bandalos & Finney, 2001), which have not been demonstrated in previous studies. Thus, previous studies are not informative about the fit between the scores from MSCEIT and the underlying model, and the proposed factor structure has been challenged (Keele & Bell, 2008). Alternative scales might therefore turn out to be more appropriate.

Fourth, Føllesdal and Hagtvet (2009) used Generalizability Theory (GT, Brennan, 2001; Cronbach, Gleser, Nanda, & Rajaratnam, 1972) and found that only some of the branches in the MSCEIT provide scores that are reliable and interpretable. They suggested some alternative scales that provide reliable and interpretable scores, which will be used in the present study in

order to assess the relationship with transformational leadership. In the following we will look further at the validity of the scores from each branch in the MSCEIT and why it is necessary with some alternative scales for Perceiving Emotions and Using Emotions.

4.1. Branch 1: Perceiving Emotions

In MSCEIT, Perceiving Emotion is measured by two tasks, *Faces* and *Pictures*, consisting of pictures of faces, abstract art, and landscapes, and one shall rate the extent to which they express various emotions. Føllesdal and Hagtvet (2009) found that the overall score for this branch contains large sources of variance not accounted for by the theory, e.g., complex combinations between type of picture, emotions expressed, and emotions assessed. Contrary to the underlying theory, the scores from Faces and Pictures seem to rank order people differently and should not be combined into an overall score. The scores are also multidimensional, i.e., while the factors in Pictures were difficult to interpret meaningfully, the scores from Faces constitute three moderately correlated factors assessing emotions that are *not* expressed by the faces. The few items that assess the particular emotions expressed by the faces are regularly excluded by the publisher before scoring the test. The remaining items apparently assess something different than proposed by the theory.

This finding is in line with other studies reporting unexpected findings for this branch. For instance, the two tasks load very differently on this branch (Fiori & Antonakis, 2011) and the scores are unrelated to other established measures of emotion perception ability (Austin, 2010; Roberts et al., 2006). The scores from Perceiving Emotions predicted leadership emergence in only one of two very similar studies (Côté, Lopes, Salovey, & Miners, 2010). Other studies have found that ability to recognize emotions expressed in faces, measured by a different test (the Diagnostic Analysis of Nonverbal Accuracy; DANVA), is positively related to transformational leadership, even after controlling for personality traits (Rubin, Munz, & Bommer, 2005).

The use of an overall score for the branch Perceiving Emotions is not supported by psychometric findings. Only the scores from Faces, but not Pictures, seem to be interpretable. Therefore, in the present study, the three alternative scales from Faces will be used, and they will be referred to by the acronyms APE/NE (Absence of Positive Emotions in faces expressing Negative Emotions), ANE/NE (Absence of Negative Emotion in faces that mostly express Negative Emotion), and ANE/PE (Absence of Negative Emotion in faces mostly expressing Positive Emotion). Though these scales seem to assess something very different from what is stated by the theory, they may nevertheless measure some important aspect of EI. For example, both ability to identify correctly that a person *is* experiencing happiness, and ability to identify that this person is *not experiencing* anger or sadness, may be important in social interaction, and hence be related to transformational leadership. We have limited knowledge about these constructs, but as these scores might measure aspects of EI, we will assess to what extent the scores from the three subscales of Faces predict transformational leadership.

4.2. Branch 2: Facilitating Thought

In MSCEIT, this branch is measured by two tasks, *Facilitation* and *Sensations*. Føllesdal and Hagtvet (2009) estimated a generalizability (reliability) coefficient of only .37 for the overall score, and found that the scores from the two tasks rank order people very differently and should not be combined into one overall score. The scores from the task Facilitation, however, were reliable, with an estimated generalizability coefficient of .62.

This finding is in line with other studies which question the validity of the scores from this branch, as the scores do not seem to measure the underlying construct adequately (Palmer et al., 2005; Rossen et al., 2008), and mixed results have been reported regarding predictive validity. For instance, contrary to expectations, the scores have been found to be unrelated to various coping styles and depressive affect (Goldenberg, Matheson, & Mantler, 2006), and the scores predicted leadership emergence in only one of two very similar studies (Côté et al., 2010). As only the task Facilitation in this branch provides generalizable scores, we will assess to what extent the scores from Facilitation may predict transformational leadership.

4.3. Branch 3: Understanding Emotions

In MSCEIT, this branch is measured by two tasks, *Changes* and *Blends*. Føllesdal and Hagtvet (2009) found that the estimated generalizability coefficient for these scores were .50, which is not very high for a 32 item scale. Other studies, however, have found that the branch scores relate in expected ways to many other variables, such as crystallized ability (Farrelly & Austin, 2007) and leadership emergence (Côté et al., 2010). As this branch provides scores that are generalizable and interpretable, they will be used in the present study. One previous study found that the scores from this branch were unrelated to transformational leadership as measured by the Multifactor Leadership Questionnaire (MLQ 5X). However, we will assess to what extent the scores will predict transformational leadership.

4.4. Branch 4: Managing Emotions

In MSCEIT, Managing Emotions is measured by two tasks: *Emotion Management* and *Emotional Relations*. Føllesdal and Hagtvet (2009) estimated a generalizability coefficient of .46, which suggests that the scores from this branch are difficult to generalize to the intended domain, and may therefore be difficult to use as a measure of EI. This is also in line with other studies reporting low reliability, e.g., one study reported a split-half reliability of only .45 for Managing Emotions (Lopes, Salovey, & Straus, 2003, p. 647). Empirical findings have also been mixed, e.g., the scores predict the perceived quality of social interactions (Lopes et al., 2003), but neither leader emergence (Côté et al., 2010) nor transformational leadership (Weinberger, 2009). However, despite

these shortcomings, this scale will be included in the present study in order to compare the findings with previous research. Thus, we will assess to what extent the scores from Managing Emotions may predict transformational leadership.

4.5. The total EI score from MSCEIT

The Total EI score in MSCEIT is the average score from the four branch scores, and is often used as an indicator of overall EI. As mentioned, several studies have questioned the validity of the branch scores from the MSCEIT (Palmer et al., 2005; Rossen et al., 2008) and some studies therefore use the Total EI score instead of the branch scores (e.g., Rossen & Kranzler, 2009). It has been questioned, however, whether an overall factor fits the data (Fiori & Antonakis, 2011; Rossen et al., 2008). It is also very difficult to interpret the Total EI score when it is based on branch scores that are unreliable and difficult to interpret. In order to compare findings with previous research, however, we will assess to what extent the Total EI score may predict transformational leadership after controlling for the FFM and GMA.

5. The need for multilevel analyses

Assessments of transformational leadership are very often based on subordinates' ratings of their leader. These ratings are not independent, and therefore violate the assumption of independency in observations underlying most statistical analyses (Muthén, 1994). That is, subordinates' ratings may be influenced by several common sources, e.g., that they rate the same leader, share the same colleagues, and share the same organizational culture.

When the multilevel nature of data are not taken into account, the estimated parameters and obtained factor structure of scores may be biased (Dyer, Hanges, & Hall, 2005). Studies assessing the relationship between transformational leadership and individual differences, however, seldom take such dependencies into account. Ignoring such dependencies in the analysis may lead us to draw incorrect conclusions about the role of individual differences in leadership. Multilevel analysis, however, takes such dependencies into account (Heck & Thomas, 2009), and the present study will therefore use a multilevel framework in order to assess the relationship between transformational leadership and EI.

6. The hypotheses tested in the present study

A major purpose of the present study is to use a multilevel framework to assess whether EI as ability can predict transformational leadership when controlling for the FFM and GMA. First, in order to be able to compare the results with previous research, we will test the predictive validity of the standard scales from the MSCEIT (i.e., the scores from the four branches and the Total EI score), despite that the scores are not supported by psychometric evidence. Second, we will test the predictive validity of the scores from a set of alternative scales from the MSCEIT, which provide more reliable and interpretable scores. Thus, in the present study, we will test the following three hypotheses:

H1. The scores from the four standard subscales from the MSCEIT predict transformational leadership after controlling for the FFM and GMA.

H2. The Total EI score from the MSCEIT predict transformational leadership after controlling for the FFM and GMA.

H3. The scores from the alternative set of scales from the MSCEIT predict transformational leadership after controlling for the FFM and GMA.

7. Method

7.1. Subjects and procedure

A total of 111 leaders provided data for this study. One group ($n = 66$) were recruited through the Administrative Research Foundation (AFF) at the Norwegian School of Economics and Business Administration (NHH) in Norway, and another group ($n = 45$) through a Norwegian municipality. In return for participation, the leaders were provided with an extensive feedback on results from a larger set of questionnaires assessing personality, EI, and leadership behavior. The leaders were instructed to select five subordinates among those who knew them well. To prevent selection bias, they were instructed to choose (among the subordinates who knew them well) the five subordinates whose last name appears first in the alphabet. The MSCEIT and the personality measure in this study were completed by a total of 111 leaders (73 men and 38 women, mean age 43 years). Among these, subordinate ratings on the questionnaires used in the present study were obtained for 104 of the executives (68 men and 36 women). They were rated by 459 subordinates (54% women, 46% men, mean age = 43.8 years), and each leader was rated by an average of 4.4 subordinates. The subordinates had on average worked under their leader for 39.5 months (SD = 38 months). All subordinates completed the questionnaires anonymously.

7.2. Measures

7.2.1. EI

The MSCEIT 2.0 (Mayer et al., 2002) was used to measure EI. The test has been translated to Norwegian by the first author and back-translated to English by a bi-lingual psychologist. It has been reviewed by the publisher and accepted as equivalent in meaning to the English version (for a further discussion of equivalence with the English version, see Føllesdal & Hagtvet, 2009). As mentioned, generalizability coefficients for the scores in the present sample was estimated in a previous study (Føllesdal & Hagtvet, 2009) and only some of the scales provided reliable and interpretable scores, i.e., APE/NE: ability to identify absent positive emotions in sad faces (five items); ANE/NE: ability to identify absent negative emotions in sad faces (six items); ANE/PE: ability to identify the absence of negative emotions in happy faces (four items); Facilitation (15 items[1]); and Understanding Emotions (32 items). Estimated generalizability coefficients for the scores from these scales were .73, .79, .73, .62, and .50, respectively, in the present sample (Føllesdal & Hagtvet, 2009).

The set of standard scales from the MSCEIT (i.e., the four branch scores and the Total EI score), were also used in the present study in order to compare the results with results from other studies. Scores and reliability were calculated as specified in the User's Manual, i.e., scores from 19 of the 141 items were excluded before calculating scale scores and Split-half reliability. The Split-half reliability coefficients were .93, .67, .51, .66, and .89 for Perceiving Emotions, Using Emotions, Understanding Emotions, Managing Emotions, and the Total EI score, respectively.

7.2.2. Transformational leadership

Subordinates rated their leader on 16 items from the official Norwegian translation of the MLQ 5X (Bass & Avolio, 2000; Hetland & Sandal, 2003), which measures the four components of transformational leadership behavior, i.e. Idealized Influence, Inspirational Motivation, Intellectual Stimulation, and Individualized Consideration. We used subordinates' ratings in order to reduce the impact of common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). One sub-dimension of Idealized Influence, relating to subordinates' attributions of leader charisma, was not included in this study, as the focus of this study was leader *behavior*, not attributions. Each scale was measured by four items, and Cronbach's alpha for the total score was .91.[2] All items focus on leadership style with a general frame of reference, rather than referring to any particular level (e.g., individual, group, or organization), an important distinction that has been discussed recently (Schriesheim, Wu, & Scandura, 2009).

7.2.3. The FFM personality factors

The official Norwegian translation of the 240-item NEO PI-R (Costa & McCrae, 1992) was used to measure the personality factors in the FFM and scored according to Norwegian norms (Martinsen, Nordvik, & Østbø, 2003).[3] In the present sample, Cronbach's alphas were .92, .90, .88, .88, and .87 for the scores from Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness, respectively.[2]

7.2.4. GMA

GMA was not measured in the present study, but in order to control for GMA, we obtained correlations between GMA and the other variables from large scale meta-analyses (see Table 1[4]). The correlation matrix was extended with these correlations, and regression analyses were run both with and without GMA as a predictor, in order to see if the regression coefficients for EI changed. The stability of the estimated regression parameters were assessed in a two-level Monte Carlo study (the procedure is described in more detail in the Results section).

8. Results

8.1. The multilevel model

A two-level model was specified due to the nested nature of the data (Heck & Thomas, 2009). That is, because several subordinates rate the same leader, the ratings are not independent. If such dependencies are not taken into account, the standard errors may be underestimated, increasing the probability for a Type I error (Byrne, 2012; Heck & Thomas, 2009). The data were analyzed with the computer program Mplus 7.0 (Muthén & Muthén, 1998–2012) using a robust maximum likelihood estimator (MLR). Ratings of 104 leaders were used, based on observations from 433 of 459 subordinates, due to missing data either on the independent variables, or on all variables except the independent variables.

---

[1] Scores from two items in Facilitation are regularly excluded by the publisher, but they were included in the set of alternative scales, as they have been carefully selected to represent the construct of interest, and there is no apparent conceptual reason to exclude them from the scale.

[2] Reliability was estimated with the Cronbach's alpha formula. Strictly speaking, this formula is inappropriate for assessing reliability for multidimensional measures, as the items represent a multifaceted measurement design (Cronbach, 2004). Cronbach's alpha is usually used, however, when estimating reliability for these scales, and therefore used in the present study.

[3] The NEO PI-R was scored by Hilmar Nordvik, Norwegian University of Science and Technology, on license from Psychological Assessment Resources, Inc.

[4] Regarding the correlation between GMA and transformational leadership: One meta-analysis reported a correlation of .16 between these constructs, based on a small number of studies presumably not corrected for range-restriction (DeRue et al., 2011). A larger meta-analysis (Judge, Colbert, & Ilies, 2004), however, reported that GMA correlated .33 with objective leader effectiveness, and .25 with perceived emergence (corrected for unreliability and range-restriction). We chose to use the value $r = .25$ in the analyses, but we also run the analyses with $r = .16$ to see if the results were different. The results, however, were highly similar and we therefore report only the results for analyses using $r = .25$ between GMA and transformational leadership.

Table 1
Correlations among variables on the between group level (lower diagonal) and within group level (upper diagonal).

	Mean	Var.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. TL	37.84	8.05		.16**	.08															
2. Sex	1.35	0.23	.03		−.02															
3. Age	43.15	67.95	.01	−.03																
4. Group	1.41	0.24	.15	.21*	.55**															
5. N	44.04	49.96	−.13	.11	−.05	−.11														
6. E	54.60	65.38	.23*	.17	−.25**	−.13	−.07													
7. O	51.25	71.67	.35**	.21*	.02	.15	.05	.58**												
8. A	52.01	103.05	.21	.31**	.44**	.43**	−.09	−.04	.16											
9. C	55.10	84.35	.07	.10	.07	−.11	−.38**	.21	−.02	.23*										
10. Total EI	48.46	26.54	.29*	.25**	.19	.22*	−.07	.10	.26**	.27*	.02									
11. Perceive	50.46	102.48	.19	.29	.20*	.28**	−.03	−.03	.16	.29**	−.06	.94**								
12. Use	46.20	46.57	.26	.22*	.28**	.31**	−.14	.07	.35**	.33**	−.02	.99**	.76**							
13. Understand	55.83	41.39	.20	.07	−.08	−.13	.13	.20	.13	−.03	.02	.74**	.28*	.19						
14. Manage	41.37	26.23	.33*	.13	.10	.09	−.19	.23*	.23	.15	.23	.87**	.52**	.63**	.38**					
15. APE/NE	64.98	321.27	.32*	.13	.16	.14	.10	−.09	.06	.15	−.15	.80**	.87**	.65**	.37*	.40**				
16. ANE/NE	37.76	179.59	.05	.27**	.18	.42**	−.03	−.10	.13	.27*	.00	.70**	.84**	.66**	.07	.33**	.73**			
17. ANE/PE	69.58	434.46	−.09	.10	.11	.16	−.10	.02	−.16	.07	.05	.64**	.71**	.45**	.13	.55**	.25	.37**		
18. Facilitation	42.6	63.62	.04	.06	.28**	.32**	−.19	−.07	.18	.24*	.08	.80**	.55**	1.00**	−.02	.69**	.47**	.52**	.44**	
19. GMA	–	41.3	.25[a]	.00	−.05[a]	–	−.09[a]	.08[a]	.33[a]	.01[a]	−.04[a]	.32[a]	.10[a]	.18[a]	.39[a]	.16[a]	.10[a]	.10[a]	.10[a]	.18[a]
Reliability				–	–	–	.92[b]	.90[b]	.88[b]	.88[b]	.87[b]	.89[b]	.89[c]	.67[c]	.73[c]	.66[c]	.73[d]	.79[d]	.73[d]	.62[d]

Note. Means, variances, and reliability for scores from the variables on the between group level are reported in the lower diagonal. Correlations between scores from the three variables on the within group level are reported in the upper diagonal. The independent variables were specified as latent variables, by using variance$_x$ × (1 − reliability$_x$). In order to obtain more precise estimates with more decimals, transformational leadership was up-scaled by 10 and the EI variables were upcaled by 100. TL = Transformational Leadership, N = Neuroticism, E = Extraversion, O = Openness to Experience, A = Agreeableness, C = Conscientiousness, Perceive = Perceiving Emotions, Using = Using Emotions, Understand = Understanding Emotions, Manage = Managing Emotions, APE/NE = Absent positive emotions in faces expressing negative emotions, ANE/NE = Absent negative emotions in faces expressing negative emotions, ANE/PE = Absent negative emotions in faces expressing positive emotions, GMA = general mental ability.

[a] The correlations specified for GMA were obtained from various meta-analyses in order to control for GMA: transformational leadership (DeRue et al., 2011; Judge et al., 2004; see explanation in footnote 4), the FFM (O'Boyle et al., 2011), the MSCEIT Total EI score (O'Boyle et al., 2011), the four branches from MSCEIT (Joseph & Newman, 2010). The correlations with GMA for the alternative scales from MSCEIT were specified as for the respective Branch scale. The correlation between GMA and age was specified to be slightly negative (Sundet, Barlaug, & Torjussen, 2004). The variance of GMA was specified to 41.34, based on information provided in a meta-analysis on GMA and leader effectiveness (Judge et al., 2004).

[b] Cronbach's alpha.

[c] Reliability was estimated using the Split-half procedure recommended in the User's Manual (Mayer et al., 2002).

[d] Reliability was estimated with Generalizability Theory (Cronbach et al., 1972), which takes into consideration the multifaceted measurement design of MSCEIT.

[*] $p < .05$.

[**] $p < .01$.

The scores from the 16 transformational leadership items were averaged into one transformational leadership score on the within-group level and one transformational leadership score on the between-group level. The score on the within-group level accounts for differences between subordinates' ratings of the leader on transformational leadership, while the score on the between-group level accounts for differences between leaders in transformational leadership, controlling for differences in ratings among subordinates. The intraclass correlation indexes the degree of correspondence within groups, and should approach zero when data are independent (Julian, 2001). That is, if the ICC is small, the leaders' scores on transformational leadership are only slightly different from each other, and no multi-level analyses are necessary. The intraclass correlation for transformational leadership, however, was .26, which is substantial, indicating that 26% of the variance in transformational leadership is due to differences between leaders. Thus, the substantial intraclass correlation indicates the need for a multilevel analysis (Muthén, 1994). On the within-group level, transformational leadership was regressed onto subordinates' sex and age. On the between group level transformational leadership was regressed onto the FFM, EI, and the control variables.

The independent variables on the between-group level (i.e., five scores from the FFM and the scores from the various scales in the MSCEIT) were corrected for measurement error by specifying the residual variance for each variable x to variance$_x$ × (1 − reliability$_x$). The estimated correlation matrix for the variables on both the within- and between-group level (after correcting for measurement error in the independent variables) is presented in Table 1.

8.2. The predictive validity of the MSCEIT scores after controlling for the FFM and GMA

We wanted to assess to what extent EI can predict transformational leadership after controlling for FFM and GMA. The zero-order correlations in Table 1 show that transformational leadership was significantly related to two of the personality factors in FFM, i.e., Extraversion ($r = .23$, $p = .023$) and Openness ($r = .35$, $p = .001$), and three of the EI scales, i.e., Managing Emotions ($r = .33$, $p = .025$), the Total EI score ($r = .29$, $p = .039$), and APE/NE ($r = .32$, $p = .047$). In order to assess the incremental validity of the EI scores over FFM and GMA, we specified different models including different sets of EI scales as predictors, in four steps.

First, in order to assess to what extent demographic variables and the FFM explained variance in transformational leadership, a model was specified where transformational leadership was regressed onto leader age, leader sex, group (representing the two contexts from which the leaders were recruited), and the FFM on the between-group level. The variables group and leader sex, however, caused estimation problems, likely due to their dichotomous values. The regression coefficients for these two variables were small and insignificant (i.e., β = −.11, $p = .532$ and β = −.10, $p = .538$, respectively), and a Wald test where the two coefficients were jointly constrained to zero was insignificant [$\chi^2(2, N = 433) = .776$, $p = .678$]. Due to estimation problems (and in order to reduce the number of parameters to be estimated) the two variables were excluded from the model. The remaining variables (i.e., the FFM and leader age) jointly explained 17.2% of the variance ($R^2 = .172$, $p = .053$) in transformational leadership, though none of the individual regression coefficients were significant. A Wald test, however, where the regression coefficients for the fives scales in FFM were jointly set to zero, suggested that the five scales in FFM as a set significantly predicted transformational leadership when controlling for leader age [$\chi^2(5, N = 433) = 13.136$, $p = .022$]. The model with the FFM and leader age as predictors will in the following be referred to as Model 0.

Second, in order to assess incremental validity for the EI scores, above the FFM and leader age (Model 0), the scores from the three sets of EI scales were included as predictors in three separate models, corresponding to the three hypotheses: Model 1: the four standard scales of EI, Model 2: the Total EI scale, and Model 3: the five alternative EI scales identified by Føllesdal and Hagtvet (2009). The generic model is displayed in Fig. 1.

Third, we wanted to assess the incremental validity of the EI scales, but there is no straightforward way in Mplus to test the significance of a difference in R^2 in two-level models (L. Muthén, personal communication, November 8, 2012). A likelihood ratio test, however, may be used (Hox, 2012; Pierce, 2012), and we therefore used a loglikelihood difference test, following the procedure recommended for the MLR estimator (Muthén, 2012, November 26). The model with the EI variables (unconstrained model) was compared to a model where the regression coefficients for the EI variables were constrained to zero, and the residual variance for transformational leadership was constrained to the value in the unconstrained model. A significant result indicates that including the EI variables as predictors explain significantly more variance in transformational leadership.

Finally, we wanted to assess whether including GMA as a predictor would affect the predictive validity of the EI variables. In order to control for GMA, a two-level Monte Carlo study was run for each model, using the Monte Carlo facility in Mplus 7.0. (We are grateful to Dr. John Antonakis, for providing the idea of using a Monte Carlo study in order to control for GMA.) For each model, the unstandardized regression coefficients were used as population values on the within-group level. As population values on the between-group level, we used the unstandardized regression coefficients obtained from a regression analysis on the between-group level matrix, including GMA as a predictor. That is, the covariance matrix on the between-group level was extended with empirically derived population values representing covariances with GMA. These values were obtained from recent meta-analyses assessing the correlation between GMA on the one hand, and the FFM, EI, and transformational leadership on the other (the correlations and their sources are presented in Table 1). A regression analysis was then run on this covariance matrix, including GMA as a predictor, and the resulting unstandardized regression coefficients were used as population values on the between-group level in a two-level Monte Carlo study. The average regression coefficients were estimated across 10,000 replications, and the stability, magnitude, and power of the estimates were assessed, in order to see whether including the GMA as a predictor affected the regression coefficients for the EI variables.
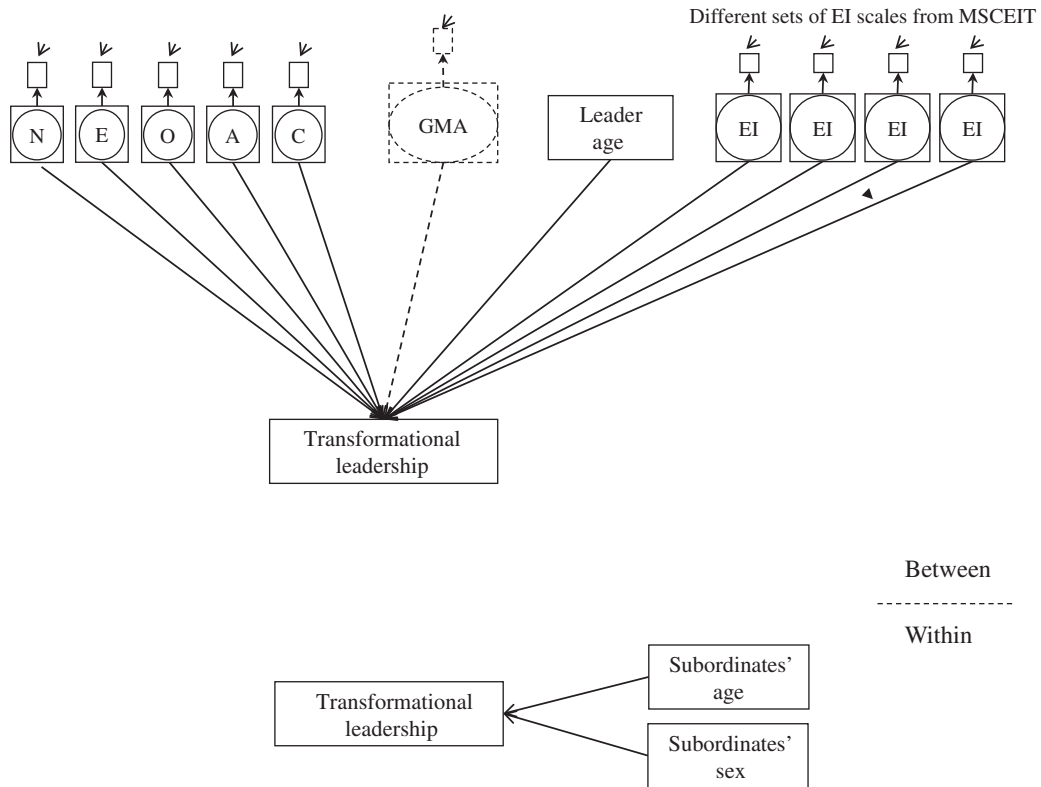
Fig. 1. Multilevel model of transformational leadership regressed on EI, the FFM, GMA, and leader age.

8.2.1. Model 1: the four standard scales in the MSCEIT as predictors

In Model 1, we assessed the extent to which the four standard scale scores from the MSCEIT could predict transformational leadership when controlling for leader age, the FFM, and GMA. Model 1 was specified similarly to Model 0, but in addition, scores from the four standard scales from the MSCEIT were included as predictors of transformational leadership on the between-group level (Table 2). The

Table 2

Model 1: Predicting transformational leadership with the four standard branch scores of EI, controlling for leader age, the FFM, and GMA.

	Model without GMA			Monte Carlo with GMA						
	β	B	p^a	β	Pop.	Est.	SD	S. E. av.	95% cov.	Power
Within level										
Subordinate age	.075	0.036	.186	.075	0.036	0.036	.025	.024	.940	.325
Subordinate sex	.171[**]	1.686	.002	.171	1.686	1.680	.505	.493	.944	.923
R^2	.034		.069							
Between level										
Leader age	−.098	−0.034	.528	−.089	−0.031	−0.032	.045	.041	.919	.155
Neuroticism	−.130	−0.055	.514	−.109	−0.046	−0.047	.054	.049	.924	.186
Extraversion	−.029	−0.011	.873	.001	0.000	0.001	.048	.044	.925	.075
Openness to experience	.298	0.107	.159	.252	0.091	0.091	.046	.042	.920	.572
Agreeableness	.198	0.059	.254	.199	0.060	0.060	.038	.035	.924	.410
Conscientiousness	−.064	−0.021	.709	−.060	−0.020	−0.020	.042	.039	.924	.106
Perceiving Emotions	−.005	−0.001	.986	.008	0.002	0.002	.038	.034	.922	.079
Using Emotions	−.066	−0.034	.879	−.076	−0.039	−0.040	.064	.060	.926	.127
Understanding Emotions	.107	0.067	.643	.063	0.039	0.039	.079	.073	.922	.111
Managing Emotions	.234	0.155	.437	.240	0.158	0.160	.084	.078	.922	.535
GMA				0.043	0.043	.057	.052	.924	.160	
R^2	.229		.062	.235						

Note. Unstandardized betas (*B*) are used as population values (Pop.) in the Monte Carlo studies and the corresponding standardized betas (β) are provided to facilitate interpretation. Est. = the average of the parameter estimates across 10,000 replications. SD = estimated standard deviation of parameter estimate across 10,000 replications. S.E. av. = the average of the estimated standard errors across 10,000 replications. 95% cov. = the proportion of the 10,000 replications for which the 95% confidence interval contains the population parameter value. Power = an estimate of power, i.e., the probability of rejecting the null hypothesis when it is false (Muthen & Muthen, 2002; Muthén & Muthén, 1998–2012).

[a] p-values are reported for the standardized regression coefficients.

[**] $p < .01$.

four standard EI scales, the FFM, and leader age explained 23% of the variance in transformational leadership ($R^2 = .229$, $p = .062$) on the between-group level. Neither the R^2 nor any of the regression coefficients were significant. Model 1 apparently explained more variance in transformational leadership than Model 0 ($R^2 = .229$ and .172, respectively), but a loglikelihood difference test indicated that the models were not significantly different [$\chi^2(5) = 3.055$, $p = .692$]. Thus, the four standard scale scores from the MSCEIT do not predict additional variance in transformational leadership when controlling for leader age and FFM.

In order to assess whether the regression coefficients for the EI variables would change when controlling for GMA, a two-level Monte Carlo study was run, including GMA as a predictor on the between-group level, following the procedure as described above. The results are displayed in the rightmost columns in Table 2. The column "Pop." displays the unstandardized regression coefficients, which were used as population values in the Monte Carlo analysis. All 10,000 replications converged, and the estimated regression coefficients (Est.) across samples were highly similar to the population values (i.e., parameter bias and standard error bias below 10% for all parameters; Muthen & Muthen, 2002). The estimated parameters were also highly similar to the estimated regression coefficients in the model without GMA. This indicates that controlling for GMA does not affect the results in any substantial way.

The estimated power coefficients for the various predictors are displayed in the rightmost column in Table 1. Among the EI scales, Managing Emotions was the strongest predictor ($\beta = .24$) with a power of .54, which means that this predictor was significant ($p < .05$) in only 54% of the 10,000 replications. Thus, a much larger sample is needed in order to detect a significant relationship between any of the EI variables and transformational leadership. In order to assess whether the four EI scales as a set could predict transformational leadership (after controlling for leader age, FFM, and GMA) a Wald test was conducted in each of the 10,000 replications, constraining the regression coefficients for the four EI variables to zero. The Wald test was significant ($p < .05$) in only 48.7% of the replications, supporting the finding that the four standard scales from the MSCEIT do not predict transformational leadership when controlling for leader age, the FFM, and GMA. Thus H1 was not supported.

8.2.2. Model 2: the Total EI score from the MSCEIT as a predictor

We wanted to assess whether the Total EI score from the MSCEIT could predict transformational leadership when controlling for leader age, the FFM, and GMA. Model 2 was therefore specified similarly to Model 0, but the Total EI score was included as a predictor on the between-group level (Table 3). The Total EI score, leader age, and the FFM explained about 20% of the variance in transformational leadership ($R^2 = .203$, $p = .045$), though none of the regression coefficients were significant. A loglikelihood difference test indicated that Model 2 did not explain more variance in transformational leadership than Model 0 [$\chi^2(2) = 2.097$, $p = .350$].

In order to see whether the inclusion of GMA as a predictor would change these results, a two-level Monte Carlo study was run, including GMA as a predictor on the between-group level (the rightmost columns in Table 3). All 10,000 replications converged, and the average estimated regression coefficients were highly similar to the population values, and highly similar to the regression coefficients in the model without GMA. The power for the Total EI score was only .28, suggesting that a large sample is needed in order to detect a significant relationship with transformational leadership. Thus, H2 was not supported, i.e., the Total EI score did not predict transformational leadership after controlling for leader age, FFM, and GMA.

8.2.3. Model 3: the five alternative scales from the MSCEIT as predictors

Finally, we wanted to assess whether the scores from the set of five alternative EI scales identified by Føllesdal and Hagtvet (2009) could predict transformational leadership when controlling for the FFM and GMA. Model 3 was specified as Model 0, but in addition, the

Table 3
Model 2: Predicting transformational leadership with the Total EI score, controlling for leader age, the FFM, and GMA.

	Model without GMA			Monte Carlo with GMA						
	β	B	p^a	β	Pop.	Est.	SD	S. E. av.	95% cov.	Power
Within level										
Subordinate age	.074	0.036		.074	0.036	0.036	.025	.024	.942	.321
Subordinate sex	.165**	1.621		.165	1.621	1.630	.502	.493	.940	.900
R^2	.032									
Between level										
Leader age	−.106	−0.037	.505	−.091	−0.032	−0.032	.044	.041	.924	.152
Neuroticism	−.117	−0.049	.535	−.105	−0.044	−0.044	.054	.050	.923	.170
Extraversion	.021	0.008	.897	.046	0.017	0.017	.047	.044	.928	.086
Openness to experience	.275	0.099	.115	.231	0.083	0.084	.045	.042	.930	.511
Agreeableness	.159	0.048	.361	.168	0.051	0.051	.038	.035	.924	.325
Conscientiousness	−.008	−0.003	.959	−.008	−0.003	−0.003	.042	.039	.930	.070
Total EI score	.185	0.109	.207	.156	0.092	0.091	.074	.069	.928	.280
GMA				.105	0.047	0.046	.056	.052	.927	.165
R^2	.203*		.045	.212						

Note. Unstandardized betas (B) are used as population values (Pop.) in the Monte Carlo studies and the corresponding standardized betas (β) are provided to facilitate interpretation. Est. = the average of the parameter estimates across 10,000 replications. SD = estimated standard deviation of parameter estimate across 10,000 replications. S.E. av. = the average of the estimated standard errors across 10,000 replications. 95% cov. = the proportion of the 10,000 replications for which the 95% confidence interval contains the population parameter value. Power = an estimate of power, i.e., the probability of rejecting the null hypothesis when it is false (Muthen & Muthen, 2002; Muthén & Muthén, 1998–2012).
[a] The p-values are for the standardized regression coefficients.
* $p < .05$.
** $p < .01$.

Table 4

Model 3: Predicting transformational leadership with the alternative scales from MSCEIT, controlling for leader age, the FFM, and GMA.

	Without GMA			Monte Carlo with GMA						
	β	B	p^a	β	Pop.	Est.	SD	S. E. av.	95% cov.	Power
Within level										
Subordinate age	.095	0.046	.089	.095	0.046	0.046	.025	.024	.938	.490
Subordinate sex	.173**	1.711	.001	.173	1.711	1.719	.505	.487	.938	.930
R^2	.038*									
Between level										
Leader age	−.141	−0.049	.392	−.118	−0.041	−0.041	.041	.037	.921	.233
Neuroticism	−.222	−0.094	.224	−.187	−0.079	−0.079	.049	.045	.920	.429
Extraversion	−.100	−0.037	.641	−.033	−0.012	−0.012	.043	.040	.921	.092
Openness to experience	.524*	0.189	.031	.427	0.154	0.154	.042	.038	.920	.964
Agreeableness	.207	0.062	.221	.214	0.064	0.064	.035	.032	.922	.520
Conscientiousness	.126	0.042	.542	.141	0.047	0.047	.039	.036	.920	.291
APE/NE	.950*	0.178	.034	1.001	0.187	0.187	.022	.020	.918	1.000
ANE/NE	−.636	−0.153	.078	−.651	−0.157	−0.156	.028	.026	.918	.999
ANE/PE	.106	0.017	.595	.087	0.014	0.014	.019	.017	.917	.170
Facilitation	−.283	−0.130	.322	−.310	−0.142	−0.142	.053	.049	.921	.801
Understanding Emotions	−.164	−0.102	.585	−.265	−0.165	−0.164	.073	.066	.918	.685
GMA				.208	0.093	0.093	.051	.047	.921	.509
R^2	.469*		.036	.496						

Note. Unstandardized betas (B) are used as population values (Pop.) in the Monte Carlo studies and the corresponding standardized betas (β) are provided to facilitate interpretation. Est. = the average of the parameter estimates across 10,000 replications. SD = estimated standard deviation of parameter estimate across 10,000 replications. S.E. av. = the average of the estimated standard errors across 10,000 replications. 95% cov. = the proportion of the 10,000 replications for which the 95% confidence interval contains the population parameter value. Power = an estimate of power, i.e., the probability of rejecting the null hypothesis when it is false (Muthen & Muthen, 2002; Muthén & Muthén, 1998–2012).

^{a} The p-values are for the standardized regression coefficients.

* $p < .05$.

** $p < .01$.

five alternative scales of EI (APE/NE, ANE/NE, ANE/PE, Facilitation, and Understanding Emotions) were included as predictors of transformational leadership on the between-group level (Table 4). The five alternative scales of EI, the FFM, and leader age explained about 47% of the variance in transformational leadership ($R^2 = .469$, $p = .036$). This relatively large R^2 appears to be due to a suppression effect: While the zero-order correlation between APE/NE and transformational leadership was moderate ($r = .32$, $p = .047$), APE/NE became a strong predictor of transformational leadership ($β = .950$, $p = .045$) once controlling for ANE/NE ($β = −.636$, $p = .093$). Moreover, Openness to experience was only moderately correlated with transformational leadership ($r = .35$, $p = .001$), but when controlling for the other predictors, it became a strong predictor ($β = .524$, $p = .035$). Further analyses revealed that this suppression effect only occurred when both APE/NE and ANE/NE were included as predictors.[5] The unexpected size and direction of the various regression coefficients might be a sign of multicollinearity, as the scores from APE/NE and ANE/NE are highly correlated ($r = .73$, $p = .000$, corrected for measurement error). The tolerance value for APE/NE and ANE/NE, however, was .28, which is acceptable, as values below .10 might indicate multicollinearity (Cohen, Cohen, West, & Aiken, 2003; Tabachnick & Fidell, 2013). Moreover, multicollinearity has little effect and may be ignored when we are interested in prediction, as is the case in the present study (Cohen et al., 2003; Tabachnick & Fidell, 2013). Model 3 apparently explained more variance in transformational leadership than Model 0 ($R^2 = .469$ and .172, respectively) and a loglikelihood difference test revealed that this difference was significant [$χ^2(6) = 21.662$, $p = .001$]. The set of alternative EI scales explained an additional 30% of the total variance in transformational leadership, after controlling for leader age and FFM.

The rightmost columns in Table 4 shows the results from a two-level Monte Carlo study after including GMA as a predictor on the between-group level. All 10,000 replications converged and the average regression coefficients were highly similar to the population values. Controlling for GMA did not change the regression coefficients in any substantial way, except somewhat larger regression coefficients for Facilitation and Understanding Emotions. The power was relatively high for four of the EI scales (i.e., over .80 for APE/NE, ANE/NE, and Facilitation; and .69 for Understanding Emotion), which indicates a relatively high probability for obtaining significant regression coefficients. In line with these results, the corresponding t-values (Est./SD) for these four scales were significant ($p < .05$), indicating that these scales from the MSCEIT were significant predictors of transformational leadership when controlling for leader age, FFM and GMA. Thus, H3 was supported.

9. Discussion

The aim of the present study was to use multilevel analyses to assess whether EI as ability, measured by the MSCEIT, can predict transformational leadership after controlling for the FFM and GMA. The results revealed that neither the four branch scores (H1) nor the

---

[5] In order to assess the suppression further, a regression analysis was run only including the alternative scales from Faces as predictors, along with leader age. A similar pattern of suppression occurred, with the following regression coefficients for the predictors: APE/NE ($β = .596$, $p = .034$), ANE/NE ($β = −.341$, $p = .246$), ANE/PE ($β = −.114$, $p = .482$), and leader age ($β = −.036$, $p = .789$), with $R^2 = .183$, $p = .240$.

Total EI score (H2) from the MSCEIT predicted transformational leadership when controlling for leader age, FFM and GMA. The set of alternative scales (H3) from the MSCEIT, however, predicted additional 30% of the total variance in transformational leadership. This finding seems to be due to a suppressor situation among two subscales from Perceiving Emotions, i.e., in the regression analysis, the scores from one subscale seem to remove irrelevant variance in the scores from the other subscale.

9.1. The predictive validity of the scores from the standard scales from the MSCEIT

The most important finding in this study was that H1 and H2 were not supported, i.e., the scores from the four branch scales and the Total EI score from the MSCEIT did not predict transformational leadership after controlling for leader age, the FFM, and GMA.

Some may question why some of the regression coefficients are relatively large but insignificant, e.g., Managing Emotions ($\beta = .234$) and Openness to Experience ($\beta = .298$) in Model 1 (Table 2). If a regular single level regression analysis had been conducted (rather than a multilevel analysis) the coefficients might have been significant, because the standard errors would have been underestimated. That is, "standard statistical tests lean heavily on the assumption of independence of the observations. If this assumption is violated (and in multilevel data this is almost always the case) the estimates of the standard errors of conventional statistical tests are much too small, and this results in many spuriously 'significant' results" (Bliese & Hanges, 2004; Hox, 2002, p. 5, for a further discussion of this point, see Bliese & Hanges, 2004; Muthén, 1997). The present data has a multilevel structure, as revealed by the relatively large ICC of .26, and a multilevel approach is therefore appropriate in order to obtain trustworthy estimates (Muthén, 1994). In order to obtain sufficient power, however, the Monte Carlo study revealed that a much larger sample is needed than the 104 leaders in the present study, due to the weak relationship between EI and transformational leadership.

Some may also question why none of the scales in FFM were significant predictors of transformational leadership. For instance, both Extraversion and Agreeableness were significantly correlated with transformational leadership (Table 1). But in the regression analysis in Model 1, where all scales in the FFM were included as predictors along with leader age, the regression coefficients and the R^2 were insignificant. This occurs because some of the predictors are correlated with each other but weakly correlated with the dependent variable (Cohen et al., 2003). That is, leader age was uncorrelated with transformational leadership ($r = .01$, *ns*), but significantly correlated with both Agreeableness ($r = .44$, $p < .01$) and Extraversion ($r = -.25$, $p < .01$). Including leader age as a predictor along with the FFM thus reduces the significance of both the R^2 and the regression coefficients. The FFM, however, is nevertheless an important predictor of transformational leadership, as FFM and leader age together explained about 17% of the variance in this variable, and the Wald test showed that the regression coefficients for these five scales were significantly different from zero.

The finding that the four branch scores and the total EI score is unrelated to subordinates' ratings of transformational leadership is in line with other findings in the literature (Harms & Credé, 2010; Weinberger, 2009). These findings may therefore indicate that EI as ability is unimportant in transformational leadership, or that the MSCEIT does not measure EI appropriately. Several researchers, however, have argued that EI as ability is important in transformational leadership, and as mentioned, when EI as ability has been measured by other means than the MSCEIT, positive results have been reported (Harms & Credé, 2010; Rubin et al., 2005). Overall, the evidence suggests that the validity of the scores from the standard scales from the MSCEIT may be questioned.

9.2. The predictive validity of the scores from the alternative set of scales from the MSCEIT

H3 concerned the predictive validity of the set of alternative scales from MSCEIT. Only one of the five scales (APE/NE) provided scores that correlated positively with transformational leadership, suggesting that ability to correctly assess that particular positive emotions are not expressed in sad faces might be important. The results, however, revealed an unexpected finding: A suppression situation occurred between this scale and one of the other alternative scales from the task Faces in Perceiving Emotions. That is, the scores measuring ability to correctly assess that particular positive emotions are not expressed in sad faces (APE/NE) became a strong predictor of transformational leadership when controlling for ability to correctly assess that particular negative emotions are not expressed in faces expressing negative emotions (ANE/NE).

The interpretation and implication of this unexpected finding is not clear. One interpretation is that APE/NE measures an ability that is important in transformational leadership, and that the measurement of this ability is purified when removing the variance shared with ANE/NE. This finding might make sense, as employees with leaders scoring high on transformational leadership tend to experience more positive emotions (Bono, Foldes, Vinson, & Muros, 2007), and positive emotions are associated with several positive outcomes, including well-being and work performance (Lyubomirsky, King, & Diener, 2005). To the extent that ability to facilitate positive emotions among subordinates is important in transformational leadership, ability to detect that subordinates are not experiencing such emotions might also be important.

On the other hand, the construct validity of the scores from these alternative scales should be assessed more thoroughly before giving emphasis to the above interpretation. Though the scales provide both reliable and interpretable scores, the present theory of EI does not account for the constructs measured. For instance, the theory of EI states that the scores from Faces should be unidimensional (Mayer et al., 2003), while the scores from Faces have been found to be multidimensional (Føllesdal & Hagtvet, 2009). Moreover, the items in Faces seem to measure something different than what is stated by the authors of the MSCEIT. That is, they state that "…the items [in the task Faces] address emotion expressed in a face […]" (Mayer, Salovey, & Caruso, 2012, p. 405). However, as mentioned, scores from five of the items in Faces are removed before scoring the test (G. Sitarenios, personal communication, December 11th, 2003), and these items are the only items representing emotions expressed in the faces (Føllesdal & Hagtvet, 2009). The scales therefore seem to measure ability to assess emotions that are not expressed by the faces. One needs a theory that can account for these alternative constructs, and evidence supporting the construct validity of the scores, in order to provide a clear interpretation of the suppressor situation in this study.

Moreover, suppressor situations are less replicable than non-suppressor situations, and need to be carefully replicated (Tzelgov & Henik, 1991). One should therefore not give too much emphasis to this finding before it has been replicated and given a more thorough theoretical explanation. We need to examine further the validity of the scores from the MSCEIT, as most of the scales provided scores that were uncorrelated with transformational leadership, and none of the standard scales did predict transformational leadership.

9.3. The validity of the scores from the MSCEIT may be questioned

There are many reasons to question the validity of the scores from the MSCEIT. One reason is that the correctness of each response in the MSCEIT is determined by the consensus method: Each item response is weighted according to the proportion of a larger sample that has chosen that particular response, and the weight indicates the correctness of that response (Mayer et al., 2002). This scoring method provides low increments in EI scores when answering difficult questions correctly, and large increments in scores when answering easy questions correctly. This may prevent one from identifying people that are highly emotionally intelligent, hence the scores may not be used to identify leaders that are highly emotionally intelligent.

A second reason to question the validity of the scores from the MSCEIT is that the operationalization does not seem optimal. As mentioned, several items in the MSCEIT provide scores that correlate weakly or negatively with the respective task scores, despite all items have been considered appropriate indicators of the same construct. Thus, "the scores are based on 122 out of 141 items because psychometric analyses on the normative sample suggested exclusion of 19 items. These were not deleted from the actual test so as to preserve a balanced layout with the same number of items for all questions" (Lopes et al., 2004, p. 1021). The scores are excluded on empirical rather than conceptual grounds, as "the tasks themselves have been selected over a decade of study, which included consideration of the conceptual connection of each task to the theory" (Mayer et al., 2004a, 2004b, p. 203). As the theory does not specify that any item shall provide negative item-total score correlations, the negative correlations may either indicate that the constructs are more complex than assumed, or that the operationalization is poor. The set of items providing positive item-scale correlations also seems to vary among samples (Lopes et al., 2003, p. 1021), suggesting that the operationalization of the constructs is weak. In the present study, we followed the publisher's regular procedure to exclude the scores from 19 of the items before calculating the total score for each branch and the Total EI score, in order to assess the predictive validity of the standard set of scales from the MSCEIT. As mentioned, for Perceiving Emotions, we excluded five items from task Faces and three items from the task Pictures. The scores from the five items that are excluded from Faces apparently assess ability to identify emotions that are present in the faces (see Føllesdal & Hagtvet, 2009), and the remaining items apparently assess emotions that are not expressed by the faces. The branch Perceiving Emotions provide multidimensional scores that are difficult to interpret, which are unrelated to scores from established measures of emotion perception (Roberts et al., 2006). This may therefore suggest that the scores from this branch are not valid. As items in the other branches in the MSCEIT also provide scores that correlate negatively with each other, one may question the operationalization of each branch.

A third reason to question the validity of the scores is that there is little evidence that the scores from the standard scales in the MSCEIT provide a good fit with the four-branch ability model. As mentioned, studies reporting good fit indices from Confirmatory Factor Analysis (CFA) are not informative about the fit between the item scores and the theoretical model, and the scales provide scores that are multidimensional (Føllesdal & Hagtvet, 2009; Keele & Bell, 2008). A recent analysis of the item scores from the MSCEIT concluded that "there does not presently appear to be compelling empirical support for the idea that EI can be conceptualized as a group of distinct but related abilities. The four-branch model of EI remains an intriguing theory without clear empirical support." (Maul, 2012, p. 516). Hence, one may question the correspondence between the scores from the scales from the MSCEIT and the underlying theory, as the scores are multidimensional, difficult to interpret, and lack reliability.

Taken together, these issues might explain why the standard scores from the MSCEIT do not predict transformational leadership, and why the alternative scales do not predict transformational leadership in any straightforward way. There are some studies, however, that have found support for expected relationships between EI as ability and leadership. For instance, the scores from Understanding Emotions and Using Emotions have been found to predict leadership emergence in small groups, even after controlling for GMA and the FFM (Côté et al., 2010). Due to the many issues related to the validity of the scores from the MSCEIT, however, future research should try to develop more refined measures of EI in order to assess the relationship with transformational leadership. It is important to assess carefully the content validity and structural validity of the scores before proceeding with studies assessing external validity.

9.4. Limitations

Some limitations of the present study may be mentioned. First, results from the Monte Carlo studies suggest that given the weak relationships between EI and transformational leadership, the present sample is too small to reveal significant relationships. However, had EI been a stronger predictor of EI, the sample size ($N = 104$) would likely be sufficient.

Second, many of the scales from the MSCEIT provide scores with low generalizability, though the Split-half coefficients are satisfactory. When estimating reliability for scores from a multi-facet measurement design like the MSCEIT, generalizability theory is more appropriate than using Split-half, as the latter may inflate reliability estimates. In the future, one may try to develop better measures of the EI constructs in order to assess whether they may predict transformational leadership.

Third, in order for EI to be an important construct in leadership, it is important to demonstrate incremental validity over both the FFM and GMA (see e.g., Antonakis et al., 2009). GMA was not measured in the present study, but correlations between GMA and other variables were obtained from meta-analyses, and the stability of the regression coefficients assessed by several Monte

Carlo studies. Future studies, however, should include measures of crystallized and fluid ability, in order to assess the incremental validity of EI measures more directly.

It is unlikely that the findings in this study are due to peculiarities in the sample, cultural differences, or poor data quality. The data were gathered in a real-world setting, and the leaders were provided feedback on individual results, and would therefore likely complete the questionnaires thoroughly. The Norwegian version of the MSCEIT have been back-translated to English and considered equivalent in meaning to the International version, and Norwegian norms has been made. Moreover, psychometric analyses of the scores demonstrate that it is the same set of items in the Norwegian and the International version of the MSCEIT that provide scores that correlate negatively with the branch scores (Føllesdal & Hagtvet, 2009), supporting the equivalence of the two language forms. The leaders' subordinates were also in a good position to assess their leader's transformational leadership behavior, as they on average had worked for their leader for several years.

10. Conclusion

The results from the present study suggest that the scores from the MSCEIT by and large are unrelated to transformational leadership after controlling for the FFM and GMA. On psychometric grounds, the validity of the scores from the MSCEIT may be questioned, and the finding that the scores do not predict transformational leadership, support this conclusion. The unexpected suppressor effect identified should be replicated and given a more thorough theoretical explanation before considered further. Future research would benefit from firmly establishing the content and structural validity of any measure of EI before assessing the relationship with transformational leadership.

References

Antonakis, J., Ashkanasy, N. M., & Dasborough, M. T. (2009). Does leadership need emotional intelligence? *The Leadership Quarterly*, *20*, 247–261, http://dx.doi.org/10.1016/j.leaqua.2009.01.006.

Ashkanasy, N. M., Hartel, C. E. J., & Daus, C. S. (2002). Diversity and emotion: The new frontiers in organizational behavior research. *Journal of Management*, *28*, 307–338, http://dx.doi.org/10.1016/S0149-2063(02)00130-7.

Ashkanasy, N. M., & Tse, B. (2000). Transformational leadership as management of emotion: A conceptual review. In N. M. Ashkanasy, & C. E. Haertel (Eds.), *Emotions in the workplace: Research, theory, and practice*. Westport, CT: Quorum Books/Greenwood Publishing Group.

Austin, E. J. (2010). Measurement of ability emotional intelligence: Results for two new tests. *British Journal of Psychology*, *101*, 563–578, http://dx.doi.org/10.1348/000712609x474370.

Bandalos, D. L., & Finney, S. J. (2001). Item parceling issues in structural equation modeling. In G. A. Marcoulides, & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bar-On, R. (1997). *BarOn Emotional Quotient Inventory (EQ-i): Technical manual*. Toronto, Canada: Multi-Health Systems.

Bar-On, R. (2000). Emotional and social intelligence. Insights from the Emotional Quotient Inventory. In R. Bar-On, & J. D. A. Parker (Eds.), *The handbook of emotional intelligence. Theory, development, assessment, and application at home, school, and in the workplace*. San Francisco: Jossey-Bass.

Bass, B. M. (1997). *Transformational leadership: Industrial, military, and educational impact*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bass, B. M. (2002). Cognitive, social, and emotional intelligence of transformational leaders. In R. E. Riggio, S. E. Murphy, & F. J. Pirozzolo (Eds.), *Multiple intelligences and leadership* (pp. 105–118). Mahwah, NJ: Lawrence Erlbaum Publishers.

Bass, B. M., & Avolio, B. (2000). *MLQ: Multifactor Leadership Questionnaire* (2nd ed.)Redwood City, CA: Mind Garden.

Bass, B. M., & Riggio, R. E. (2006). *Transformational leadership* (2nd ed.)Mahwah, NJ: Lawrence Erlbaum Associates.

Bliese, P. D., & Hanges, P. J. (2004). Being both too liberal and too conservative: The perils of treating grouped data as though they were independent. *Organizational Research Methods*, *7*, 400–417, http://dx.doi.org/10.1177/1094428104268542.

Bono, J. E., Foldes, H. J., Vinson, G., & Muros, J. P. (2007). Workplace emotions: The role of supervision and leadership. *Journal of Applied Psychology*, *92*, 1357–1367, http://dx.doi.org/10.1037/0021-9010.92.5.1357.

Bono, J. E., & Judge, T. A. (2004). Personality and transformational and transactional leadership: A meta-analysis. *Journal of Applied Psychology*, *89*, 901–910, http://dx.doi.org/10.1037/0021-9010.89.5.901.

Boyatzis, R. E., Goleman, D., & Rhee, K. S. (2000). Clustering competence in emotional intelligence: Insights from the Emotional Competence Inventory. In R. BarOn, & J. D. A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace*. San Francisco, CA: Jossey-Bass.

Brennan, R. L. (2001). *Generalizability theory*. New York: Springer.

Brown, F. W., & Moshavi, D. (2005). Transformational leadership and emotional intelligence: A potential pathway for an increased understanding of interpersonal influence. *Journal of Organizational Behavior*, *26*, 867–871, http://dx.doi.org/10.1002/Job.334.

Byrne, B. (2012). *Structural equation modeling with Mplus*. New York: Routledge.

Caruso, D. R., Mayer, J. D., & Salovey, P. (2002). Emotional intelligence and emotional leadership. In R. E. Riggio, S. E. Murphy, & F. J. Pirozzolo (Eds.), *Multiple intelligences and leadership* (pp. 55–74). Mahwah, NJ: Lawrence Erlbaum Associates.

Caruso, D. R., & Salovey, P. (2004). *The emotionally intelligent manager*. San Francisco, CA: Jossey Bass.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R. Professional manual*. Lutz, FL: Psychological Assessment Resources.

Côté, S., Lopes, P. N., Salovey, P., & Miners, C. T. H. (2010). Emotional intelligence and leadership emergence in small groups. *The Leadership Quarterly*, *21*, 496–508, http://dx.doi.org/10.1016/j.leaqua.2010.03.012.

Cronbach, L. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, *64*, 391–418, http://dx.doi.org/10.1177/0013164404266386.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurement: Theory of generalizability for scores and profiles*. New York, NY: Wiley.

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, http://dx.doi.org/10.1016/S0191-8869(03)00240-X.

DeRue, D. S., Nahrgang, J. D., Wellman, N., & Humphrey, S. E. (2011). Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. *Personnel Psychology*, *64*, 7–52, http://dx.doi.org/10.1111/j.1744-6570.2010.01201.x.

Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly*, *16*, 149–167, http://dx.doi.org/10.1016/j.leaqua.2004.09.009.

Farrelly, D., & Austin, E. J. (2007). Ability EI as an intelligence? Associations of the MSCEIT with performance on emotion processing and social tasks and with cognitive ability. *Cognition and Emotion*, *21*, 1043–1063, http://dx.doi.org/10.1080/02699930601069404.

Fiori, M., & Antonakis, J. (2011). The ability model of emotional intelligence: Searching for valid measures. *Personality and Individual Differences*, *50*, 329–334, http://dx.doi.org/10.1016/j.paid.2010.10.010.

Føllesdal, H., & Hagtvet, K. A. (2009). Emotional intelligence: The MSCEIT from the perspective of generalizability theory. *Intelligence*, *37*, 94–105, http://dx.doi.org/10.1016/j.intell.2008.08.005.

George, J. M. (2000). Emotions and leadership: The role of emotional intelligence. *Human Relations, 53*, 1027–1055.

Goldenberg, I., Matheson, K., & Mantler, J. (2006). The assessment of emotional intelligence: A comparison of performance-based and self-report methodologies. *Journal of Personality Assessment*, *86*, 33–45, http://dx.doi.org/10.1207/s15327752jpa8601_05.

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, http://dx.doi.org/10.1177/1548051809350894.

Heck, R. H., & Thomas, S. L. (2009). An introduction to multilevel modeling techniques. NY: Routledge.

Hetland, H., & Sandal, G. M. (2003). Transformational leadership in Norway: Outcomes and personality correlates. *European Journal of Work and Organizational Psychology*, *12*, 147–170, http://dx.doi.org/10.1080/13594320344000057.

Hox, J. J. (2002). Multilevel analysis. Mahwah, New Jersey: Lawrence Earlbaum Associates.

Hox, J. J. (November 9th). Re: [MULTILEVEL] Testing R-square difference in two-level models [Electronic mailing list message]. Retrieved from. https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=multilevel

Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology*, *95*, 54–78, http://dx.doi.org/10.1037/a0017286.

Judge, T. A., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, *89*, 542–552, http://dx.doi.org/10.1037/0021-9010.89.3.542.

Julian, M. W. (2001). The consequences of ignoring multilevel data structures in nonhierarchical covariance modeling. *Structural Equation Modeling*, *8*, 325–352, http://dx.doi.org/10.1207/s15328007sem0803_1.

Keele, S. M., & Bell, R. C. (2008). The factorial validity of emotional intelligence: An unresolved issue. *Personality and Individual Differences*, *44*, 487–500, http://dx.doi.org/10.1016/j.paid.2007.09.013.

Kim, S., & Hagtvet, K. A. (2003). The impact of misspecified item parceling on representing latent variables in covariance structure modeling: A simulation study. *Structural Equation Modeling*, *10*, 101–127, http://dx.doi.org/10.1207/S15328007sem1001_5.

Kupers, W., & Weibler, J. (2006). How emotional is transformational leadership really? Some suggestions for a necessary extension. *Leadership and Organization Development Journal*, *27*, 368–383, http://dx.doi.org/10.1108/01437730610677972.

Lindebaum, D., & Cartwright, S. (2010). A critical examination of the relationship between emotional intelligence and transformational leadership. *Journal of Management Studies*, *47*, 1317–1342, http://dx.doi.org/10.1111/j.1467-6486.2010.00933.x.

Lopes, P. N., Brackett, M. A., Nezlek, J. B., Schütz, A., Sellin, I., & Salovey, P. (2004). Emotional intelligence and social interaction. *Personality and Social Psychology Bulletin*, *30*, 1018–1034, http://dx.doi.org/10.1177/0146167204264762.

Lopes, P. N., Salovey, P., & Straus, R. (2003). Emotional intelligence, personality, and the perceived quality of social relationships. *Personality and Individual Differences*, *35*, 641–659, http://dx.doi.org/10.1016/S0191-8869(02)00242-8.

Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin*, *131*, 803–855, http://dx.doi.org/10.1037/0033-2909.131.6.803.

Martinsen, Ø., Nordvik, H., & Østbø, L. E. (2003). Norsk utgave av Revised NEO Personality Inventory (NEO PI-R) [Norwegian version of the Revised NEO Personality Inventory (NEO PI-R)]. Oslo, Norway: Gyldendal Akademisk.

Maul, A. (2012). Examining the structure of emotional intelligence at the item level: New perspectives, new conclusions. *Cognition and Emotion*, *26*, 503–520, http://dx.doi.org/10.1080/02699931.2011.588690.

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, *27*, 267–298, http://dx.doi.org/10.1016/S0160-2896(99)00016-1.

Mayer, J. D., Roberts, R. D., & Barsade, S. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology*, *59*, 507–536, http://dx.doi.org/10.1146/annurev.psych.59.103006.093646.

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

Mayer, J. D., Salovey, P., & Caruso, D. R. (2002). Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT). User's manual. Toronto, Canada: Multi-Health Systems.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004a). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*, *15*, 197–215, http://dx.doi.org/10.1207/s15327965pli1503_02.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004b). A further consideration of the issues of emotional intelligence. *Psychological Inquiry*, *15*, 249–255, http://dx.doi.org/10.1207/s15327965pli1503_05.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence: New ability or eclectic traits? *American Psychologist*, *63*, 503–517, http://dx.doi.org/10.1037/0003-066X.63.6.503.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2012). The validity of the MSCEIT: Additional analyses and evidence. *Emotion Review*, *4*, 403–408, http://dx.doi.org/10.1177/1754073912445815.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion*, *3*, 97–105, http://dx.doi.org/10.1037/1528-3542.3.1.97.

Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods and Research*, *22*, 376–398.

Muthén, B. O. (1997). Latent variable modeling of longitudinal and multilevel data. In A. Raftery (Ed.), Sociological methodology (pp. 453–480). Boston, MA: Blackwell Publishers.

Muthén, B. O. (November 26). Difference testing using the loglikelihood. Retrieved 2012, November 26, from. http://www.statmodel.com/chidiff.shtml

Muthén, B. O., & Muthén, L. K. (1998–2012). Mplus user's guide (6th ed.).

Muthen, L. K., & Muthen, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, *9*, 599–620, http://dx.doi.org/10.1207/s15328007sem0904_8.

O'Boyle, E. H., Humphrey, R. H., Pollack, J. M., Hawver, T. H., & Story, P. A. (2011). The relation between emotional intelligence and job performance: A meta-analysis. *Journal of Organizational Behavior*, *32*, 788–818, http://dx.doi.org/10.1002/job.714.

Palmer, B. R., Gignac, G., Manocha, R., & Stough, C. (2005). A psychometric evaluation of the Mayer–Salovey–Caruso Emotional Intelligence Test Version 2.0. *Intelligence*, *33*, 285–305, http://dx.doi.org/10.1016/j.intell.2004.11.003.

Petrides, K. V. (2010). Trait emotional intelligence theory. *Industrial and Organizational Psychology*, *3*, 136–139, http://dx.doi.org/10.1111/j.1754-9434.2010.01213.x.

Pierce, S. J. (November 9). Re: [MULTILEVEL] Testing R-square difference in two-level models [Electronic mailing list message]. Retrieved from. https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=multilevel

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, http://dx.doi.org/10.1037/0021-9010.88.5.879.

Roberts, R. D., Schulze, R., O'Brien, K., MacCann, C., Reid, J., & Maul, A. (2006). Exploring the validity of the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT) with established emotions measures. *Emotion*, *6*, 663–669, http://dx.doi.org/10.1037/1528-3542.6.4.663.

Rode, J. C., Mooney, C. H., Arthaud-day, M. L., Near, J. P., Rubin, R. S., Baldwin, T. T., et al. (2008). An examination of the structural, discriminant, nomological, and incremental predictive validity of the MSCEIT V2.0. *Intelligence*, *36*, 289–382, http://dx.doi.org/10.1016/j.intell.2007.07.002.

Rosete, D., & Ciarrochi, J. (2005). Emotional intelligence and its relationship to workplace performance outcomes of leadership effectiveness. *Leadership and Organization Development Journal*, *26*, 388–399, http://dx.doi.org/10.1108/01437730510607871.

Rossen, E., & Kranzler, J. H. (2009). Incremental validity of the Mayer–Salovey–Caruso Emotional Intelligence Test Version 2.0 (MSCEIT) after controlling for personality and intelligence. *Journal of Research in Personality*, *43*, 60–65, http://dx.doi.org/10.1016/j.jrp.2008.12.002.

Rossen, E., Kranzler, J. H., & Algina, J. (2008). Confirmatory factor analysis of the Mayer–Salovey–Caruso Emotional Intelligence Test V 2.0 (MSCEIT). *Personality and Individual Differences, 44*, 1258–1269, http://dx.doi.org/10.1016/j.paid.2007.11.020.

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, http://dx.doi.org/10.5465/amj.2005.18803926.

Salovey, P., & Mayer, J. D. (2002). The positive psychology of emotional intelligence. In C. R. Snyder, & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 159–171). New York: Oxford University Press.

Salovey, P., & Sluyter, D. J. (Eds.). (1997). *Emotional development and emotional intelligence: Educational implications.* New York: Basic Books, Inc.

Schriesheim, C. A., Wu, J. B., & Scandura, T. A. (2009). A meso measure? Examination of the levels of analysis of the Multifactor Leadership Questionnaire (MLQ). *The Leadership Quarterly, 20*, 604–616, http://dx.doi.org/10.1016/j.leaqua.2009.04.005.

Sundet, J. M., Barlaug, D. G., & Torjussen, T. M. (2004). The end of the Flynn effect?: A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence, 32*, 349–362, http://dx.doi.org/10.1016/j.intell.2004.06.004.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.)New Jersey: Pearson.

Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications, and applications. *Psychological Bulletin, 109*, 524–536, http://dx.doi.org/10.1037/0033-2909.109.3.524.

Walter, F., Cole, M. S., & Humphrey, R. H. (2011). Emotional intelligence: Sine Qua Non of Leadership or Folderol? *Academy of Management Perspectives, 25*, 45–59, http://dx.doi.org/10.5465/amp.2011.59198449.

Weinberger, L. A. (2009). Emotional intelligence, leadership style, and perceived leadership effectiveness. *Advances in Developing Human Resources, 11*, 747–772, http://dx.doi.org/10.1177/1523422309360811.