

## Emotional intelligence: The hype, the hope, the evidence

2015

Emotion Researcher

The Official Newsletter of the International Society for Research on Emotion

John Antonakis, Professor of Organizational Behavior  
University of Lausanne

In 2001, I arrived at the psychology department of Yale University to undertake a postdoc; I thought exciting things were going on at the time in terms of broadening our understanding of human abilities. I went there as a pilgrim—to work on leadership (and “practical intelligence”) with Robert Sternberg. Peter Salovey’s work on emotional intelligence (EI) was beginning to get a lot of traction at that time. At the outset, I thought that the ideas and research programs around these alternative notions of intelligence, and particularly in EI, were laudable and done with good intent. Yet, when I left Yale, I left as a skeptic. Why?

To be clear, let me first say that Salovey and Mayer have reshaped conventional thinking; one of their key contributions to science has been the construct of emotional intelligence (EI), defined as the “*ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions* (Salovey & Mayer, 1989-1990, p. 189). Although more than 20 years have passed since the publication of this seminal piece, the scientific credentials of EI are still very much in question. There are three major issues concerning the EI construct:

1. There is great disagreement in the scientific community with respect the conceptualization of EI, whether EI should be measured as an ability, akin to an intelligence test (which relies on scoring subjects’ performance using “objective” scoring keys), or if EI should be measured as a trait analogous to how personality is measured (using self-reports). This debate has no end in sight; the proponents of ability measures, with whom I sympathize, say that it is folly to ask subjects to self-rate an ability (e.g., like asking a subject to self-rate their general intelligence). The self-rating aficionados argue that ability measures are bedeviled with measurement issues.
2. Very poor testing standards have been used to assess whether EI can be measured, whether it is distinct from personality and IQ, and whether it matters for outcomes like job or leadership performance or other outcomes; when rigorous testing standards are used, the effects of EI vanish (e.g., Antonakis, 2009; Antonakis & Dietz, 2011a, 2011b; Cavazotte, Moreno, & Hickmann, 2012; Fiori & Antonakis, 2011, 2012). Because there is evidence that EI is strongly correlated with personality and general intelligence, and disentangling the effects of EI from the effects of the variable with

which EI correlates is often ignored, whatever results are reported in the literature about the apparent predictive power of EI are highly biased and not trustworthy (Antonakis, Bendahan, Jacquart, & Lalive, 2010).

3. Exaggerated claims have been made, by popular writers but also publishers of what are apparently reputable EI tests, about the accuracy with which EI measures can predict success in a number of performance domains. These claims are highly problematic because many well-meaning human resources directors or other professionals use EI tests in selection or for clinical purposes; doing so is neither ethical nor economical.

I will address each of the three points in more detail below. But let me first lay my conclusion on the table: I am very doubtful that emotional intelligence as currently theorized and measured is a valid scientific construct (Antonakis, 2003, 2004, 2009); there are theoretical and empirical reasons for my positions. I return to the theoretical issues at the end of this article and focus here on the empirical ones. In order to be valid, a psychometric measure ought to have at least construct validity (it should elicit responses that reveal differences in the latent ability we are trying to measure) and incremental validity (it should predict practically useful outcomes beyond what competing constructs like personality and intelligence can already predict). But, EI as currently measured fails on both accounts. All measures of EI currently in use elicit responses meant to capture abstract knowledge about emotions, without gauging how individuals use such knowledge in real world situations. That is, the latent ability researchers are trying to measure, the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions, is not fully captured: no test exists whatsoever to predict people's *actions* in using this EI ability in real life situations, whereas this is the primary evidential ground we use to distinguish emotionally intelligent from non-emotionally intelligent people in ordinary life.

Furthermore, there are a priori reasons to think that EI as currently measured cannot be distinguished from other competing constructs, and add to their predictive power. For example, it is pretty clear that personality traits (mostly emotional stability, agreeableness, and openness) correlated strongly with measures of EI and play a role in monitoring one's own feelings and others' feelings. Moreover, the ability to use emotion knowledge to guide one's thinking and action depends in part on one's ability to learn condition-action scripts (i.e., if a subordinate is distraught by previous failure, act empathetically towards him or her). Scripts require repeated exposure to condition-action links (Abelson, 1981) and being able to abstract from these depends on general intelligence (Gottfredson, 1997, 2002; Schmidt, 2002; Schmidt & Hunter, 1998); the very definition of intelligence is about the ability to learn. Whether the scripts are about emotions or other concepts is, I believe, not relevant and neural underpinnings about there being more than one general intelligence are lacking (Antonakis, Ashkanasy, & Dasborough, 2009).

As concerns my position on the three major issues I listed above, first, I have been rather clear that if there is anything to EI—then ability tests, like the MSCEIT (Mayer-Salovey-Caruso Emotional Intelligence test) are the way to go (Antonakis, et al., 2009). It is harder to provide

socially-desirable response on an ability test, because the “correct” answer is not so obvious. In self-rated tests, however, it is all too easy to select “I agree” when presented with statements like “I am sensitive to the feelings and emotions of others” (this is an item taken from the WLEIS test of Wong & Law, 2002); additionally, these self-rated tests overlap too much with known constructs (e.g., personality, intelligence) and do not predict performance measures once when controlling for the known constructs (Joseph, Newman, & O'Boyle, in press).

Yet, evidence is starting to pile up that even the venerable MSCEIT, the Mayer-Salovey-Caruso Emotional Intelligence test, and the theory underlying it may be flawed (Amelang & Steinmayr, 2006; Antonakis & Dietz, 2010; Fiori & Antonakis, 2011, 2012; Fiori et al., 2014; Føllesdal & Hagtvet, 2009; Maul, 2012; Zeidner, Matthew, & Roberts, 2001). The problems identified by these researchers are manifold. Briefly, and focusing on the MSCEIT—the flagship measure of the ability movement—its four factors or branches (perceiving emotions, understanding emotions, managing emotions, and using emotions in thought) do not appear to reflect a general (i.e., higher-order) construct of EI, as conceived by the original authors (Mayer, Caruso, & Salovey, 1999). Such results raise doubts about its construct validity, and call into question the idea that one’s global EI is causally responsible for the scores obtained on the four branches (see Fiori & Antonakis, 2011 for further discussion).

An additional worry relates to the scoring system used in MSCEIT (Fiori & Antonakis, 2011; Legree et al., 2014; Matthews, Zeidner, & Roberts, 2002; Maul, 2012; Roberts, Zeidner, & Matthews, 2001). Unlike most IQ tests, which have clearly established objective answers, EI tests are calibrated based on experts' judgments (“expert” scoring) or majority respondent ratings (“consensus” scoring). Unfortunately for the developers of the test, correlations between scores of experts on the test correlate close to unity (i.e.,  $r \approx 1.00$ ) with scores of lay individuals; how can commoners score so highly with experts? Who are these experts anyway? It was reported that they are 21 emotion theorists gathered at ISRE 2000, but who are they and what are they experts on exactly? Are the items on the MSCEIT test too easy? Can a consensus-based scoring system truly detect expertise in a tested domain, which presumably requires being above average?

Finally, there is growing empirical evidence that the MSCEIT factors overlap way too much with personality and intelligence (e.g., Fiori & Antonakis, 2011; Legree, et al., 2014; Schulte, Ree, & Carretta, 2004)—far more than its architects are willing to acknowledge. The problem is not that correlations are present; but, that they are so strong suggests EI is in serious danger of redundancy.

A lot of EI “eureka” moments—that is apparent discoveries showing that EI really matters for success—have been based on very lax and even sloppy application of basic psychometric testing principles. The litmus test of validity is incremental validity, whereby the key construct is examined alongside competing constructs. For instance, if one claims to have a fast horse (EI), this claim should be examined by testing the horse’s speed relative to other horses known to be champions (personality, intelligence). Yet, most studies testing EI fail to control for competing constructs like IQ and the big five personality factors, and this using robust designs (e.g., correcting for measurement error, using a correct model specification, etc.).

This has not prevented developers of EI, who have commercial interests in the construct, from making extravagant claims about the predictive power of emotional intelligence. For example, Goleman, Boyatzis, and McKee (2002, p. 251)—whose trait model of EI blends almost every imaginable trait that is not general intelligence into an unwieldy mix (Sternberg, 1999)—stated the following: “To get an idea of the practical business implications of these [EI] competencies, consider an analysis of the partners’ contributions to the profits of a large accounting firm ... those with strength in the self-regulation competencies added a whopping 390 percent incremental profit—in this case, \$1,465,000 more per year. By contrast, significant strengths in analytic reasoning abilities added just 50% more profit. Thus, purely cognitive abilities help—but the EI competencies help far more.” What all this means from a statistical and validity point of view is quite unclear.

To really know, however, if a variable predicts performance one must refer to well-designed studies and ideally to meta-analyses, a statistical technique that pools together the results of many independent studies. These analyses have clearly established that the single most important predictor of work performance is general intelligence, and that the correlation between general intelligence and performance increases as job complexity increases (Salgado, Anderson, Moscoso, Bertua, & De Fruyt, 2003; Schmidt & Hunter, 1998). These results, when corrected for statistical artefacts (measurement error and restriction of range) show that the correlation of general intelligence to job performance is about .70, even when controlling for conscientiousness and emotional stability (Schmidt, Shaffer, & Oh, 2008). To get an idea of the predictive strength of a correlation of .70, we can convert this statistic to a practical measure of effect (Rosenthal & Rubin, 1982); doing so shows that 85% of individuals who are above the median on general intelligence will score above the median in job performance; however, only 15% of individuals below the median on general intelligence will score above the median in job performance. In other words, a smart individual is more than five times more likely to have above median performance than will a less-smart individual. These validity coefficients are as good as it gets in psychology; I would be very surprised to see anything beat this in my lifetime!

EI does not even come close to matching the predictive power of general intelligence with respect to job performance. As summarized by Van Rooy and Viswesvaran (2004, p. 87) “EI did not evidence incremental validity over GMA [general mental ability, i.e., general intelligence]. However, GMA did significantly predict performance beyond that explained by EI. Thus, the claims that EI can be a more important predictor than cognitive ability (e.g., Goleman, 1995) are apparently more rhetoric than fact.” When testing for incremental validity via meta-analysis, researchers usually assume certain validity coefficients for competing constructs. Researchers must use the most accurate coefficients for those competing constructs so as to obtain unbiased estimates. If they do not, they will severely tilt the research record.

For instance, some researchers have “plugged in” values of .47 for the relation between intelligence and job performance in the statistical model tested (e.g., Joseph & Newman, 2010) instead of more realistic values (which are substantially higher, as I reported above); ironically, though, even in this meta-analysis, where the cards were stacked in EI’s favor by using a validity coefficient for general intelligence that was understated by a hefty margin, these

researchers reported that “measures of ability models of EI show only a modicum of incremental validity over cognitive ability and personality traits” (p. 69); another recent meta-analysis showed too that ability EI tests are not incrementally valid over personality and intelligence (see “Stream 1” results in Table 6: O’Boyle, Humphrey, Pollack, Hawver, & Story, 2011). Interestingly too is that even for a performance domain like leadership that is heavily emotions-based, there is no meta-analytic evidence that EI matters for leadership when controlling for personality or the big five (Harms & Credé, 2010).

Why then is research still undertaken with such zeal and why do practitioners still use tests of EI while repeating the mantra that “EI matters much for performance”? It is difficult to answer this question. One certainty is that there are commercial interests at stake (i.e., selling EI tests is a big business), which makes it even more important to pay attention to all the evidence, and ensure that EI tests are viable and do what they are intended to do prior to marketing them. Unfortunately, there is no body like the FDA to regulate psychometric tests—it is up to the academic market and the good consciousness of publishers to decide.

This leads me to the third issue I want to raise about EI, namely that it is borderline unethical to sell EI tests under the false pretense that they are scientifically proven, when it is pretty clear that there is an ongoing scientific controversy on whether or not EI tests capture what they intend to capture and add anything to the predictive power of other existing constructs. Although lives are not directly at stake by using invalid EI tests, I find it unconscionable that money is being made from claims that are not backed up by very solid data.

At first blush, there is a clear distinction to be drawn in this respect between scientific researchers of EI and popularizers of EI. For example, Mayer, Salovey, and Caruso have repeatedly distanced themselves from Goleman’s wildly unsubstantiated claims about the predictive power of EI, noting: “Our own work never made such claims, and we actively critiqued them” (2008). I praise them for taking this stance. The line is blurred, however, when we read what the publisher of MSCEIT has to say about the test: “A large and growing body of independent scientific research has identified it [EI] as the single most important determinant of superior functioning; emotionally intelligent people succeed because they are better able to read and deal with social complexity. As confirmed by independent academic research, one-quarter to nearly one-half of all job success can be attributed to Emotional Intelligence.” All I can say to that is: *Gasp!* How can researchers of this caliber allow their publisher to make such claims? And how are such claims any different from the claims Goleman and company continue to peddle to the unsuspecting public?

A related worry is that the commercialization of MSCEIT stands in the way of its scientific testing. Suppose I wanted to run a test on the incremental validity of MSCEIT with 200 participants, the sort of study I have argued is sorely needed to establish whether or not MSCEIT predicts job performance beyond intelligence and personality. I would need to spend \$52.50 for a manual purchased with a researcher’s discount, and \$6 per participant to take the on-line test, making for a total of \$1,252.50. This is an amount that many researchers might not be able to afford. So the very price of MSCEIT makes it hard for the scientific community to test its

scientific credentials while the test is being sold as being backed by solid scientific evidence. Is there not something wrong with this picture?

Where do we go from here? Although I have so far focused on what's wrong with EI, I acknowledge that Mayer, Salovey and other pioneers of the research program are right that emotion-management ability is differentially distributed in the population and that it matters for various performance domains; the question is whether we need to call this ability EI and whether smarts and personality traits alone can explain why some people are better than others at understanding and managing emotions and at reaping the benefits of this ability in consequential settings. To fully answer this question, we need robust testing, using the most rigorous psychometric standards, and to pay attention to the evidence—all of it and particularly that from robust studies. We also need to move away from tests that merely capture people's knowledge of emotions, and develop tests that test the real-world side of "emotional intelligence." It is one thing to measure knowledge of emotional processes or hypothetical intents with respect to these emotional processes as is done with the MSCEIT—it is another thing entirely to enact the correct decision "on the fly." In other words, how individuals respond to items on the MSCEIT test (e.g., using emotions intelligently in decision making) does not necessarily map on how they would act in real-world settings (Fiori & Antonakis, 2012). Nothing in the current EI measures gauges this ability in high fidelity situations.

A final worry is that there is no "a priori" reason to expect that having a highly attuned "emotional radar" can help individuals take appropriate decisions, especially in emotionally-charged situations. I have called this problem the "curse of emotion" (Antonakis, et al., 2009). Briefly, it can become more difficult to make effective decisions as one becomes increasingly sensitive to one's own or others' emotional states. This is because in many circumstances the morally or economically advisable decision may require hurting the feelings of some stakeholders (i.e., "cut the branch to save the tree" type decisions), and increased sensitivity to such feelings can lead to either failure to act properly or massive costs in "emotional labor" for the decider. This suggests that future research on EI should take the decision context into account too, and not simply assume that being "attuned" to the emotions of others is advantageous in all contexts. Also, I have yet to see a test that measures, on one hand the ability to be highly attuned to emotions while, on the other hand "setting emotions aside" when needed in decision making (i.e., not being bogged down by them) or using them in appropriate "doses" in decision making (see Antonakis, et al., 2009).

To conclude, debate in science is healthy and is needed. And, there comes a time when we have to rethink theories or measurement strategies and then "move on." This time is nigh for those doing research in emotional intelligence. My message to them is: "Drop the hype, keep the hope, but pay attention to all the evidence. Not only is the moral thing to do; it is also the economical thing to do."

## References

- Abelson, R. P. (1981). Psychological status of the script concept. *American Psychologist*, 36(7), 715-729.
- Amelang, M., & Steinmayr, R. (2006). Is there a validity increment for tests of emotional intelligence in explaining the variance of performance criteria? *Intelligence*, 34(5), 459-468.
- Antonakis, J. (2003). [Why "emotional intelligence" does not predict leadership effectiveness: A comment on Prati, Douglas, Ferris, Ammeter, and Buckley](#). *The International Journal of Organizational Analysis*, 11(4), 355-361.
- Antonakis, J. (2004). [On why "emotional intelligence" will not predict leadership effectiveness beyond IQ or the "big five": An extension and rejoinder](#). *Organizational Analysis*, 12(2), 171-182.
- Antonakis, J. (2009). ["Emotional intelligence": What does it measure and does it matter for leadership?](#) In G. B. Graen (Ed.), *LMX leadership--Game-Changing Designs: Research-Based Tools* (Vol. VII, pp. 163-192). Greenwich, CT: Information Age Publishing.
- Antonakis, J., Ashkanasy, N. M., & Dasborough, M. T. (2009). [Does leadership need emotional intelligence?](#) *The Leadership Quarterly*, 20(2), 247-261.
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). [On making causal claims: A review and recommendations](#). *The Leadership Quarterly*, 21(6), 1086-1120.
- Antonakis, J., & Dietz, J. (2010). [Emotional intelligence: On definitions, neuroscience, and marshmallows](#). *Industrial and Organizational Psychology*, 3(2), 165-170.
- Antonakis, J., & Dietz, J. (2011a). [Looking for Validity or Testing It? The Perils of Stepwise Regression, Extreme-Scores Analysis, Heteroscedasticity, and Measurement Error](#). *Personality and Individual Differences*, 50(3), 409-415.
- Antonakis, J., & Dietz, J. (2011b). [More on testing for validity instead of looking for it](#). *Personality and Individual Differences*, 50(3), 418-421.
- Cavazotte, F., Moreno, V., & Hickmann, M. (2012). Effects of leader intelligence, personality and emotional intelligence on transformational leadership and managerial performance. *The Leadership Quarterly*, 23(3), 443-455.
- Fiori, M., & Antonakis, J. (2011). [The ability model of emotional intelligence: Searching for valid measures](#). *Personality and Individual Differences*, 50(3), 329-334.
- Fiori, M., & Antonakis, J. (2012). [Selective attention to emotional stimuli: What IQ and openness do, and emotional intelligence does not](#). *Intelligence*, 40(3), 245-254.
- Fiori, M., Antonietti, J. P., Mikolajczak, M., Luminet, O., Hansenne, M., & Rossier, J. (2014). What Is the Ability Emotional Intelligence Test (MSCEIT) Good for? An Evaluation Using Item Response Theory. *PLoS ONE*, 9(6).
- Føllesdal, H., & Hagtvet, K. A. (2009). Emotional intelligence: The MSCEIT from the perspective of generalizability theory. *Intelligence*, 37, 94-105.
- Goleman, D., Boyatzis, R., & McKee, A. (2002). Primal leadership: Realizing the power of emotional intelligence. *Personnel Psychology*, 55(4), 1030-1033.
- Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24(1), 79-132.
- Gottfredson, L. S. (2002). Where and why g matters: Not a mystery. *Human Performance*, 15(1/2), 25-46.

- Harms, P. D., & Credé, M. (2010). Remaining Issues in Emotional Intelligence Research: Construct Overlap, Method Artifacts, and Lack of Incremental Validity. *Industrial and Organizational Psychology, 3*(2), 154-158.
- Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology, 95*, 54-78.
- Joseph, D. L., Newman, D. A., & O'Boyle, E. H. (in press). Why Does Self-Reported Emotional Intelligence Predict Job Performance? A Meta-Analytic Investigation of Mixed EI. *Journal of Applied Psychology*, <http://dx.doi.org/10.1037/a0037681>.
- Legree, P. J., Psotka, J., Robbins, J., Roberts, R. D., Putka, D. J., & Mullins, H. M. (2014). Profile Similarity Metrics as an Alternate Framework to Score Rating-Based Tests: MSCEIT Reanalyses. *Intelligence, 47*(0), 159-174.
- Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional Intelligence: Science and Myth*. Cambridge, MA: MIT Press.
- Maul, A. (2012). The Validity of the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT) as a Measure of Emotional Intelligence. *Emotion Review, 4*(4), 394-402.
- Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 27*(4), 267-298.
- Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence - New ability or eclectic traits? *American Psychologist, 63*(6), 503-517.
- 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*(5), 788-818.
- Roberts, R. D., Zeidner, M., & Matthews, G. (2001). Does Emotional Intelligence Meet Traditional Standards for an Intelligence? Some New Data and Conclusions. *Emotion, 1*(3).
- Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74*(2), 166-169.
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., & De Fruyt, F. (2003). International validity generalization of GMA and cognitive abilities: A European community meta-analysis. *Personnel Psychology, 56*(3), 573-605.
- Salovey, P., & Mayer, J. D. (1989-1990). Emotional Intelligence. *Imagination, Cognition and Personality, 9*(3), 185-211
- Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance, 15*( 1/2).
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262-274.
- Schmidt, F. L., Shaffer, J. A., & Oh, I. S. (2008). Increased accuracy for range restriction corrections: Implications for the role of personality and general mental ability in job and training performance. *Personnel Psychology, 61*(4), 827-868.
- Schulte, M. J., Ree, M. J., & Carretta, T. R. (2004). Emotional Intelligence: Not much more than g and personality. *Personality and Individual Differences, 37*(5), 1059-1068.
- Sternberg, R. J. (1999). Review of Daniel Goleman's Working with Emotional Intelligence. *Personnel Psychology, 52*(3), 780-783.
- Van Rooy, D. L., & Viswesvaran, C. (2004). Emotional intelligence: A meta-analytic investigation of predictive validity and nomological net. *Journal of Vocational Behavior, 65*, 71-95.

- Wong, C. S., & Law, K. S. (2002). The effects of leader and follower emotional intelligence on performance and attitude: An exploratory study. *The Leadership Quarterly*, 13(3), 243-274.
- Zeidner, M., Matthews, G., & Roberts, R. D. (2001). Slow Down, You Move Too Fast: Emotional Intelligence Remains an "Elusive" Intelligence. *Emotion*, 1(3), 265-275.