

Assessment of Emotional Management ability: a Rasch Model Analysis of the brief Situational
Test of Emotional Management (STEM-B)

Carolina da Motta^{a)}, Célia Barreto Carvalho^{a), b)}, Paula Castilho^{a)}, Michele T. Pato^{c), d)}

a) Cognitive-Behavioral Research Centre (CINEICC), College of Psychology and Education Sciences,
University of Coimbra, Portugal

b) Division of Psychology, Department of Educational Sciences, University of Azores, Azores, Portugal

c) Director & Professor Institute for Genomic Health, SUNY Downstate Medical Center, Brooklyn, NY

d) Keck School of Medicine, Department of Psychiatry and Behavioral Sciences, University of Southern
California

Correspondence concerning this article should be addressed to: Carolina da Motta, CINEICC,
Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra, Rua do Colégio
Novo, Apartado 6153, 3001-802, Coimbra, Portugal. E-mail: carolina.d.motta@uac.pt

Role of funding source: This project is funded by the FCT (Portuguese Foundation for Science
and Technology, grant SFRH/BD/110308/2015, and cosponsored by the Operational Program for
the Human Potential (POPH) / European Social Fund (FSE).

Conflict of Interests: None of the authors has a conflict of interest to disclose.

Assessment of Emotional Management ability: a Rasch Model Analysis of the brief Situational Test of Emotional Management (STEM-B)

Abstract

Background: Emotional Intelligence (EI) is a thriving research field with a variety of applications extending from clinical psychology, to academic or professional accomplishment and organizational success. The role of Emotion Management as the crucial skill in achieving high-performance or subjective wellbeing has increased the demand for effective and psychometrically sound measurements of emotional management ability. **Method:** The current study aimed to explore the psychometric properties of the Portuguese version of the brief Situational Test of Emotional Management (STEM-B) using Rasch Model (Item Response Theory) in a sample of 899 participants from the general population. **Results:** After inspection of assumptions and the removal of 26 participants presenting problematic response patterns, the global fit indicated the measure had good metric properties, with most items aligning vertically across the logits scale and presenting an adequate range of item difficulty and item fit in the STEM-B. **Discussion:** Amidst an unset debate over the measurement paradigms and very placement of the EI construct and skills, the STEM-B is a psychometrically sound measure for specific testing of emotional management skills with a wide potential utility for transcultural studies, applicable in several fields and the advancement on EI ability assessment and theoretical models.

Keywords: emotional management, emotional intelligence, maximum performance, assessment, Rasch Model

Conceptual models of Emotional Intelligence (EI) typically suggest a set of components or skills aimed at adaptive intra and interpersonal functioning and management of social interaction (Bar-On, 2006; Cherniss, 2004; Mayer, Salovey, & Caruso, 2000). The current study focuses on emotional regulation or management, identified as the 4th branch of the Mayer-Salovey EI model, which consists of different behaviors or regulations strategies aimed at modifying the emotional experience or expressions (i.e. minimizing negative and maximizing positive emotional states) or facilitating goal-directed behaviors (Aldao, Nolen-Hoeksema, & Schweizer, 2010; Dixon-Gordon, Bernecker, & Christensen, 2015; Gratz & Roemer, 2004; Tamir & Ford, 2012). The Mayer-Salovey theory proposes a hierarchical model in which each branch depends on skills from lower branches, meaning that emotion management is at the top of the hierarchy, built upon skills related to perceiving emotions, using them to facilitate the thought process and understanding of emotional experience (Mayer, Roberts, & Barsade, 2008). Within a social interaction, emotional management skills can be used to regulate one's own emotions and influencing the emotional states of others (Dixon-Gordon et al., 2015), and their maladaptive or adaptiveness may be assessed based on the outcomes or consequences of the use of different emotion regulation strategies (e.g. increasing or decreasing well-being, social support, distress, psychological symptoms).

Starting from the popularization of the EI concept over the past two decades, EI, and emotional management skills in particular have received a significant attention due to its impact on significant outcomes like professional performance, teaching, healthcare, well-being, leadership, business (Austin, 2010; Mayer, Salovey, & Caruso, 2004). Regarding the latter, there has been a significant dissemination of findings on Emotional Intelligence being a decisive set of

skills in the entrepreneurial field to achieve professional success, by improving leadership and management skills, which provided additional impulse to the multiplication of programs aiming at developing the long-sought emotional management skills for professionals in leading roles in all kinds of organizations (Kelly & Kaminskienė, 2016; Martina, Denisa, & Mariana, 2015; Zhang, Cao, & Wang, 2018). But despite the popular and research interest around EI, the current debate on theoretical models and adequate methodological approach to measure EI and its component is still far from reaching a consensus or pointing directions towards how the adaptiveness or maladaptiveness of responses or strategies and their outcomes should be determined. This latter aspect is fundamental to mental health research and intervention, including the fields of psychopathology and clinical psychology, as they may exacerbate or diminish anxiety states, depressive symptoms or establishing consistent patterns that lead to personal invalidation and interpersonal problems – which are ultimately a core feature of personality disorders and one of the most detrimental outcome of any psychological disorder (Aldao et al., 2010).

The conceptual overlap and terminological interchangeability of constructs of emotion management and emotion regulation in the current literature increases the difficulty in the definition of this process, and the extrinsic and intrinsic strategies (e.g. behaviors, goal-orientation, motives) that operationalize the process. This often poses an additional obstacle to researchers attempting to systematize the current knowledge of the field, to compare or replicate previous research studies, to test the predictive assumptions of current conceptualizations or to advance with theoretical refinements of existing EI models. Nevertheless, recent evidence points to the necessity of revising the 4-branch EI model by suppressing the second branch (using emotions to facilitate other tasks or processes), as its conceptual redundancy poses significant

measurement problems (Anguiano-Carrasco, MacCann, Geiger, Seybert, & Roberts, 2015; Austin, 2010; Joseph & Newman, 2010; MacCann, Joseph, Newman, & Roberts, 2014).

The EI field currently divides itself into 2 types of EI that are seemingly unrelated constructs, trait EI and ability EI, the first relating to the personality and the second to the intelligence domain. This separation had significant implications to assessment methodology (Austin, 2010; MacCann & Roberts, 2008; Petrides, 2011). Many of the traditional or more widely disseminated tools were made available for commercial use and are presented in formats that not only can be susceptible to biases, like item transparency or response formats that may influence the relationship of EI scores and relevant variables such as personality, intelligence, academic achievement, or well-being (Anguiano-Carrasco et al., 2015; Davis & Humphrey, 2012; Joseph & Newman, 2010; Libbrecht & Lievens, 2012). These issues raise questions relating to the very placement of the construct in the general framework of multiple intelligence models (Husin, Santos, Ramos, & Nordin, 2013; MacCann et al., 2014; Pardeller, Frajo-Apor, Kemmler, & Hofer, 2017), due to possible confounding effects with verbal or other cognitive abilities (when the variance in the tests of EI cannot be solely attributed to the construct), or resulting in biases in assessment when applied into particular populations (e.g. clinical populations) or contexts (e.g. academic, high-stakes job applications, where faking or social desirability becomes more likely). For instance, the MSCEIT, one of the most disseminated and debated measures of EI, resorts to different scoring procedures when assessing different EI components (multiple-choice and rate the extent) and presented issues such as high correlation with IQ. Other measures related with trait EI tend to correlate more with personality traits and are evaluated by self-report inventories, while ability EI correlates strongly to intelligence tests and crystallized abilities, as they are typically presented as problem-solving tests (MacCann &

Roberts, 2008; Roberts, Schulze, & MacCann, n.d.). The remarkable demand for psychometrically sound tools for the assessment of EI ability has led to the investment in new assessment paradigms and tests to overcome shortcomings - mostly related to validity or consistency - observed in several studies involving more mainstream assessment instrument (Allen, Rahman, et al., 2014; Allen, Weissman, Hellwig, MacCann, & Roberts, 2014; Austin, 2010; MacCann & Roberts, 2008; Mestre, MacCann, Guil, & Roberts, 2016).

Several authors advocate the advantages of maximum performance testing for EI abilities over typical performance test (Anguiano-Carrasco et al., 2015; Austin, 2010; Libbrecht & Lievens, 2012; MacCann & Roberts, 2008). A recent trend has provided evidence for the Situational Judgement Tests (STJ) paradigm as a valid approach to the assessment of emotional abilities, an approach that was already present in some MSCEIT tests, and has been further developed by MacCann & Roberts (2008) in two new measures of EI ability: the Situational Test of Emotional Understanding and the Situational Test of Emotional Management, to which the studies by Austin (2010) and Libbercht & Lievens (2012) provided initial evidence of construct validity of those measures.

The current study focuses on one of these tests, specifically the psychometric properties of brief version of the Situational Test of Emotional Management (STEM-B). In a previous study by Allen et al (2014), an 18-item version was developed from the 44-items Situational Test of Emotional Management (STEM), indicating the shorter version preserved the characteristics of the longer version. However, investigation of this measure in independent samples is still lacking. The STEM-B is a performance test depicting several interpersonal scenarios (some taking place in professional context, others in personal contexts or without a specific context) to which respondents must choose the most efficient strategies within 4 response options. The

availability of a brief measure of emotional management skills that is easily accessible is of utmost importance to diversify the alternatives for EI ability measurement, and when working with specific populations or when time constraints is a concern for researchers, and particularly for researchers working with non-English speaking researchers, to which the options are even more scarce and delay the development of cross-national studies that contribute to the field of EI ability and related constructs. We approach the analysis with Item Response Theory (IRT), a more advantageous method in EI research when compared to classical test theory (Anguiano-Carrasco et al., 2015), in a large sample of Portuguese-speaking participants from the general population.

Method

Participants and Procedures

The first stage consisted in obtaining permission to translate and use the Situational Test of Emotion Management – Brief (described in the Measures section) from the American Psychological Association. A Portuguese bilingual psychologist translated the instructions and items, and a bilingual technician residing in the United States of America back translated the test. A senior psychologist revised the translation and found no deviations from the content of the original version of the test. Prior to its administration to a wider sample, the STEM-B was administered to 5 participants, who did not report difficulties regarding item's clarity or comprehension.

The current study is a part of a large research project registered in the National Data Protection Committee (Authorization nr. 5853/2016) and approved by the Ethical Boards of the University of Coimbra (Deliberation of January 12th, 2017) and Azores University (Deliberation

nr. 5/2016). All participants were informed about the research goals, warranting the anonymity and the voluntary character of participation, obeying to research ethical principles of scientific research including human participants, and only participants who provided their written informed consent were administered a research protocol including the STEM-B. Underage participants were contacted through local schools and a signed informed consent was obtained from their legal representatives prior to participation.

A sample of 899 adolescents and adults from the general population from Portuguese mainland and Azores islands participated in this study. Forty percent of participants were males ($n=360$) and 60% were females ($n=539$), between 14 and 72 years old ($M = 22.46$; $SD = 9.98$). Most participants were single ($N = 748$; 83.2%), followed by married/in a civil union ($N = 125$; 13.9%), 21 participants were divorced (2.3%) and 2 widowed (0.2%). Three participants (0.3%) did not report their civil status. Most participants ($N = 388$; 43.2%) had completed elementary school and 344 (38.3%) were currently attending to or completed mandatory education (high school). The remaining 151 (16.7%) participants reported having concluded higher education (college, masters or doctoral degree) or alternative curricula (e.g. attendance to professional schools), while 16 participants did not report information about their education (1.7%).

Statistical Analyses

Rasch Analysis was performed with WINSTEPS Rasch Analysis (version 3.93, SWREG Inc., 2017) and the remaining statistical procedures were carried out using IBM SPSS Statistics (version 23 for Microsoft Windows, IBM Inc. Armonk, NY).

Because the STEM-B is a multiple-choice test format, items were dichotomized into correct or incorrect responses and analysed through Rasch Model analysis (Rasch, 1961). In

Rasch Models, calculations transform the persons and item parameters to a unit “measure” (θ or theta) that is distributed along a continuum, like a ruler. Each unit of measure of θ are log-odd units or “logits”, a scale with theoretical ranges being \pm infinite, but typically ranges between an amplitude of ± 5 (Prieto & Velasco, 2006), and 0 localizes the average difficulty point set for the measure.

Infit Mean Square (Infit MNSQ) and Outfit Mean Square (Outfit MNSQ) are observed to estimate the fit of the data to the model, both for items and persons. The Infit refers to the weighted mean square, providing information about items and possible structural problems (Baker, 2001; Prieto & Velasco, 2006). According to Linacre (Linacre, 2011), Infit statistics values between 0.5 and 1.5 are productive for measurement; values greater than 2.0 can degrade the measurement; values between 1.5 and 2.0 are unproductive for measurement construction; and values smaller than 0.5 are less productive for measurement, but not degrading. The Outfit is the unweighted mean square and is useful for diagnosing outliers (Linacre, 2011).

Measure

Situational Test for Emotional Management – Brief (Allen, Rahman, et al., 2014; MacCann & Roberts, 2008). This tool retains the 18 most informative items from the 44-item version of the Situational Test for Emotional Management (STEM). The STEM was developed based on Situational Judgment Paradigm and through semistructured interviews and expert evaluations of response options. Because results may be influenced by response formats, the final multiple-choice rating format was chosen by a quasi-experimental design comparing the multiple-choice with rate-the-extent (Allen, Rahman, et al., 2014; MacCann & Roberts, 2008). This procedure aimed at separating the test and construct effects, multiple-choice format was

conveyed the more adequate due to its convergence to emotional and cognitive process more related to intelligence tests (choosing the best option), over the divergent thinking pattern required in rate-the-extent formats (simultaneously considering and selecting from equally good options).

Results

The test assumption of one-dimensionality was assessed through Principal Component Analysis (or contrasts) of standardized residuals variance. The Eigenvalue of unexplained variance obtained on the second contrast was 1.40, and because this value was inferior to 2, the test presents no multidimensionality problems (Linacre, 2011; Raïche, 2005). The assumption of local independence was confirmed from low residual correlations between items, with values ranging from $r = -.12$ to $r = .25$ (Linacre, 2011). The total raw variance explained by the scale was 25.2%.

Global fit statistics

The model tested has involved all participants and items from the STEM-B. Global fit measures for Persons and items are presented in Table 1. All items presented adequate adjustments, with average Infit and Outfit values of .99 ($SD = 0.10$) and 1.02 ($SD = .22$), respectively. The maximum value of the Outfit was of 1.57 suggests the absence of outliers or items with poor adjustment. The amplitude of the measure for the items ranged from -1.53 to 1.73 logits and measure's standard error was low (between .07 and .09, $M = 0.08$; $SD = 0.01$). Person fit showed appropriate to *Infit* ($M = 1.00$, $SD = .96$) and *Outfit* ($M = 1.02$, $SD = .46$) average values. The maximum values of outfit inform about the existence of abnormal response patterns. The inspection of extreme infit and outfit values revealed that about 26 participants (2.9%) presented response patterns diagnosed matching "careless" or "lucky guessing" (Linacre

& Wright, 1994), which would result in a poor fit to the model. As misfitting response patterns are possible cause of measure distortions, the model was reanalyzed after excluding the 26 participants presenting problematic response patterns.

Table 1.
Global fit statistic of STEM-B (N = 899)

	Reliability	Separation		Measure (θ)	Model Error	Infit	Outfit
Person fit	.63	1.30	M	.22	.58	1.00	1.02
			SD	.96	.09	.22	.46
			Max	3.17	1.05	1.86	7.52
			Min	-3.17	.51	.57	.28
Item fit	.99	11.23	M	.00	.08	.99	1.02
			SD	.90	.01	.10	.22
			Max	1.73	.09	1.11	1.57
			Min	-1.53	.07	.75	.60

Note. Max = maximum; Min = minimum.

Table 2 presents the improvements observed in global fit statistics after removing outliers, indicating the test has the necessary conditions to produce an instrument with adequate metric properties. Item's fit statistic now fall within adequate values of infit and outfit, without significant indicators of measure degradation within the model. Reliability and separation values for person was .64 and 1.33, indicating a lower ability to discriminate participants in the sample according to their level of performance, while the item reliability and separation were .99 and 11.83, suggesting a good item difficulty hierarchy. The Cronbach alpha (KR-20) for the scale was .62.

Table 2.

Global fit statistic of STEM-B (N = 873)

	Reliability	Separation		Measure (θ)	Model Error	Infit	Outfit
Person fit	.64	1.33	<i>M</i>	.27	.56	1.00	.98
			<i>SD</i>	.94	.07	.22	.33
			Max	3.21	1.05	1.72	2.27
			Min	-3.20	.52	.56	.25
Item fit	.99	11.83	<i>M</i>	.00	.08	.99	.98
			<i>SD</i>	.95	.01	.09	.16
			Max	1.90	.10	1.12	1.19
			Min	-1.68	.07	.77	.61

Note. Max = maximum; Min = minimum.

Items-person map

Figure 1 depicts the distribution of persons and items across the measure (θ). The majority of items align vertically across the logits scale. The only 2 sets of parallel items suggest a similar ability of emotional management and could be excluded, if necessary. Provided those items did not present any problems within the model and refer to distinct interpersonal scenarios theoretically relevant to the construct at hand, they were maintained.

The θ average for the items is zero by convention and the average θ for persons was .27 logits, and most participants fall within a range of 4 logits (between -2 and 2). Item 16 and 10 were the least frequently rated correct, suggesting those items can evaluate better emotional management ability than observed in most participants. The item measuring the highest degrees of emotional management are, thus, item 16 (nearer point 2 of the logits scale), and the item measuring the lesser degree was 18 (placed near point -2 of the logits scale). Finally, as presented in more detail in table 3, each item’s statistics of STEM-B present adequate values of Infit and

Outfit values, low standard errors and adequate (i.e. positive) point-biserial correlations to the total measure.

Table 3.

Total score, θ , Standard Error, Infit e Outfit by Item, and Point-bisserial correlations between item and total score of STEM-B (n = 873)

Item	Total score	θ	<i>SE</i>	Infit	Outfit	<i>r</i>
1	430	.32	.07	1.05	1.08	.33
2	480	.05	.07	1.06	1.07	.32
3	312	.97	.08	1.05	1.10	.30
4	468	.12	.07	.98	1.02	.40
5	466	.13	.07	.97	.98	.41
6	578	-.50	.08	.99	.97	.40
7	638	-.88	.08	.96	.90	.41
8	371	.64	.07	1.07	1.10	.30
9	473	.09	.07	1.10	1.18	.27
10	257	1.3	.08	1.05	1.19	.27
11	610	-.70	.08	.99	.92	.40
12	343	.80	.07	1.12	1.14	.25
13	671	-1.12	.09	.86	.79	.50
14	697	-1.32	.09	.78	.68	.56
15	341	.81	.07	1.01	1.01	.35
16	173	1.90	.09	1.04	1.05	.26
17	646	-.94	.08	1.00	.95	.38
18	737	-1.68	.1	.77	.61	.56
<i>M</i>	482.8	.00	.08	.99	.98	
<i>SD</i>	158.7	.95	.01	.09	.16	

Note. θ = measure; *SE* = standard error; *r* = point-measure correlation (between observations and θ).

Discussion

The current study aimed to test the psychometric properties of brief version of the Situational Test of Emotional Management (STEM-B) in an independent sample from the initial study by Allen et al (2014). We used Rasch Model (Item Response Theory), which allows the creation of a measurement unit from a sample’s performance that considers the item’s difficulty within the same construct, useful to ability measurement models.

Because IRT assumes the local independence of items (responses are not similar and do not depend on each other), reliability estimates may differ from more typical indices of reliability which are prone to be artificially inflated by item correlations (e.g. Cronbach's alpha, KR-20). The reliability (KR-20) obtained in the current study was lower than those from the studies of the 44-item version (Austin, 2010; MacCann & Roberts, 2008) and the brief version by Allen et al. (2014), but still within acceptable standards. The high item reliability showed the item difficulty hierarchy could be located across the latent variable with a significant precision. The lower person reliability and separation values may be a result of the reduced number of items, or the sample presenting a narrow ability range – being constituted by participants from the general population and no other possible extreme groups, for instance. This latter issue was also raised in the study by Libbrecht & Lievens (Libbrecht & Lievens, 2012) when using the longer version of the measure (STEM) to explore its relationship to other EI ability measures. Nevertheless, the separation values indicate the possibility of distinguishing 2 levels of performance (e.g. high/low performance). Similar findings were obtained in a study of the longer version of the STEU, which included a sample with higher average age and years schooling (MacCann & Roberts, 2008). For this reason, using a shorter form of the measure with a simple score procedure (correct/incorrect) can be a useful approach when working with vulnerable populations (e.g. clinical samples, in- or outpatients), when emotional management is a complimentary variable, or when researchers face assessment time constraints or participant fatigue is a concern. When such constraints are not a present, it would still be recommendable to use a longer version when the discrimination of individuals by more than 2 levels of performance or assessing performance in a wider range of scenarios may be needed.

While the multiple response format allows to circumvent faking responses or social desirability, it is important to keep in mind that current EI theories have not been sufficiently developed to guide the appraisal and choice of correct responses in this kind of tests, and the correctness of responses are mainly based on empirical or expert evaluations (psychologists, counselors, psychiatrists). The current study is, therefore, bound to this limitation. Future efforts should be developed in clarifying the algorithms to resolving complex and nuanced decisions involved in social interactions in which emotional skills management are based and manifested. The use of video vignettes or other multimedia support opens the possibility to the development of more dynamic ways to assessment of different emotional strategies involved in emotional management, as participants may be able to process verbal and non-verbal clues in a more fluid manner.

Overall, the STEM-B is a psychometrically adequate measure for specific testing of emotional management skills with a wide potential utility, as measures of this nature can not only be useful for academic or professional fields, but also in aiding the evaluation socio-emotional skills deficits and devising targeted interventions in clinical or educational settings, and devising effective training programs that prevent psychopathology and promote psychosocial adjustment and well-being. More importantly, the STEU-B is an accessible measure to researchers that provide clear scoring procedures that allow to respond to the current assessment needs in the EI ability field. The advantages of disseminating a Portuguese version of the STEU-B as a novel and cost-effective measure made available through APA for professionals working with Portuguese-speaking communities has increased utility to transcultural studies involving the high number of Portuguese native-speakers worldwide (Lewis, Simons, & Fennig, 2016) and fostering empirical and theoretical advancements on EI models and related constructs.

References

- Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review, 30*(2), 217–237.
<https://doi.org/10.1016/j.cpr.2009.11.004>
- Allen, V. D., Rahman, N., Weissman, A., MacCann, C., Lewis, C., & Roberts, R. D. (2014). The Situational Test of Emotional Management - Brief (STEM-B): Development and validation using item response theory and latent class analysis. *Personality and Individual Differences, 81*, 195–200. <https://doi.org/10.1016/j.paid.2015.01.053>
- Allen, V. D., Weissman, A., Hellwig, S., MacCann, C., & Roberts, R. D. (2014). Development of the situational test of emotional understanding – brief (STEU-B) using item response theory. *Personality and Individual Differences, 65*, 3–7.
<https://doi.org/10.1016/j.paid.2014.01.051>
- Anguiano-Carrasco, C., MacCann, C., Geiger, M., Seybert, J. M., & Roberts, R. D. (2015). Development of a Forced-Choice Measure of Typical-Performance Emotional Intelligence. *Journal of Psychoeducational Assessment, 33*(1), 83–97.
<https://doi.org/10.1177/0734282914550387>
- Austin, E. J. (2010). Measurement of ability emotional intelligence: results for two new tests. *British Journal of Psychology (London, England : 1953), 101*(Pt 3), 563–578.
<https://doi.org/10.1348/000712609X474370>
- Baker, F. B. (2001). *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation. <https://doi.org/10.1111/j.1365-2362.2010.02362.x>
- Bar-On, R. (2006). The Bar-On model of emotional-social intelligence (ESI). *Psicothema*.
- Cherniss, C. (2004). Intelligence, Emotional. *Encyclopedia of Applied Psychology, 315–319*.

<https://doi.org/10.1016/B0-12-657410-3/00565-1>

Davis, S. K., & Humphrey, N. (2012). The influence of emotional intelligence (EI) on coping and mental health in adolescence: divergent roles for trait and ability EI. *Journal of Adolescence*, 35(5), 1369–79. <https://doi.org/10.1016/j.adolescence.2012.05.007>

Dixon-Gordon, K. L., Bernecker, S. L., & Christensen, K. (2015). Recent innovations in the field of interpersonal emotion regulation. *Current Opinion in Psychology*, 3, 36–42. <https://doi.org/10.1016/j.copsyc.2015.02.001>

Gratz, K. L., & Roemer, L. (2004). Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure, and Initial Validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, 26(1), 41–54. <https://doi.org/10.1023/B:JOBA.0000007455.08539.94>

Husin, W. N. I. W., Santos, A., Ramos, H. M., & Nordin, M. S. (2013). The Place of Emotional Intelligence in the ‘Intelligence’ Taxonomy: Crystallized Intelligence or Fluid Intelligence Factor? *Procedia - Social and Behavioral Sciences*, 97, 214–223. <https://doi.org/10.1016/j.sbspro.2013.10.225>

Joseph, D. L., & Newman, D. A. (2010). Emotional Intelligence: An Integrative Meta-Analysis and Cascading Model. *Journal of Applied Psychology*. <https://doi.org/10.1037/a0017286>

Kelly, E. J., & Kaminskienė, N. (2016). Importance of emotional intelligence in negotiation and mediation. *International Comparative Jurisprudence*, 2(1), 55–60. <https://doi.org/10.1016/J.ICJ.2016.07.001>

Lewis, P., Simons, G., & Fennig, C. (2016). Summary by language size. In P. Lewis, G. Simons, & C. Fennig (Eds.), *Ethnologue: Languages of the World* (19th ed.). Dallas, Texas: SIL International. Retrieved from <http://www.ethnologue.com/statistics/size>

- Libbrecht, N., & Lievens, F. (2012). Validity evidence for the situational judgment test paradigm in emotional intelligence measurement. *International Journal of Psychology, 47*(6), 438–447. <https://doi.org/10.1080/00207594.2012.682063>
- Linacre, J. M. (2011). *A User's Guide to Winsteps/ Ministeps: Rasch-Model Computer Programs*. [https://doi.org/ISBN 0-941938-03-4](https://doi.org/ISBN%200-941938-03-4)
- Linacre, J. M., & Wright, B. D. (1994). (Dichotomous Mean-square) Chi-square fit statistics. *Rasch Measurement Transactions, 8*(2), 360. Retrieved from <https://www.rasch.org/rmt/rmt82a.htm>
- MacCann, C., Joseph, D. L., Newman, D. A., & Roberts, R. D. (2014). Emotional intelligence is a second-stratum factor of intelligence: Evidence from hierarchical and bifactor models. *Emotion, 14*(2), 153–166. <https://doi.org/10.1037/a0034755>
- MacCann, C., & Roberts, R. D. (2008). New paradigms for assessing emotional intelligence: Theory and data. *Emotion, 8*(4), 540–551. <https://doi.org/10.1037/a0012746>
- Martina, M., Denisa, M., & Mariana, S. (2015). Emotional Intelligence of Managers. *Procedia Economics and Finance, 26*, 1119–1123. [https://doi.org/10.1016/S2212-5671\(15\)00939-9](https://doi.org/10.1016/S2212-5671(15)00939-9)
- Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: emotional intelligence. *Annual Review of Psychology, 59*, 507–36. <https://doi.org/10.1146/annurev.psych.59.103006.093646>
- Mayer, J. D., Salovey, P., & Caruso, D. (2000). Models of emotional intelligence. *Handbook of Intelligence, 1*, 319–342. <https://doi.org/10.1080/14766080903069364>
- Mayer, J. D., Salovey, P., & Caruso, D. (2004). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry, 15*(3), 197–215. https://doi.org/10.1207/s15327965pli1503_02

- Mestre, J. M., MacCann, C., Guil, R., & Roberts, R. D. (2016). Models of Cognitive Ability and Emotion Can Better Inform Contemporary Emotional Intelligence Frameworks. *Emotion Review*, 8(4), 322–330. <https://doi.org/10.1177/1754073916650497>
- Pardeller, S., Frajo-Apor, B., Kemmler, G., & Hofer, A. (2017). Emotional Intelligence and cognitive abilities – associations and sex differences. *Psychology, Health & Medicine*, 22(8), 1001–1010. <https://doi.org/10.1080/13548506.2016.1255766>
- Petrides, K. V. (2011). Ability and Trait Emotional Intelligence.
- Prieto, G., & Velasco, A. D. (2006). Visualização espacial, raciocínio indutivo e rendimento acadêmico em desenho técnico. *Psicologia Escolar e Educacional*, 10(1), 11–20. <https://doi.org/10.1590/S1413-85572006000100002>
- Raîche, G. (2005). Critical Eigenvalue Sizes (Variances) in Standardized Residual Principal Components Analysis (PCA). Retrieved May 8, 2017, from <http://www.rasch.org/rmt/rmt191h.htm>
- Rasch, G. (1961). Probabilistic Models for Some Intelligence and Attainment Tests. *Information and Control*. [https://doi.org/10.1016/S0019-9958\(61\)80061-2](https://doi.org/10.1016/S0019-9958(61)80061-2)
- Roberts, R. D., Schulze, R., & MacCann, C. (n.d.). The Measurement of Emotional Intelligence: A Decade of Progress? In *The SAGE Handbook of Personality Theory and Assessment: Volume 2 — Personality Measurement and Testing* (pp. 461–482). 1 Oliver’s Yard, 55 City Road, London EC1Y 1SP United Kingdom : SAGE Publications Ltd. <https://doi.org/10.4135/9781849200479.n22>
- Tamir, M., & Ford, B. Q. (2012). When feeling bad is expected to be good: Emotion regulation and outcome expectancies in social conflicts. *Emotion*, 12(4), 807–816. <https://doi.org/10.1037/a0024443>

Zhang, L., Cao, T., & Wang, Y. (2018). The mediation role of leadership styles in integrated project collaboration: An emotional intelligence perspective. *International Journal of Project Management*, 36(2), 317–330. <https://doi.org/10.1016/J.IJPROMAN.2017.08.014>