12-19-2023

How many emotional intelligence abilities are there? An examination of four measures of emotional intelligence

John Mayer
University of New Hampshire, Durham, jack.mayer@unh.edu

David R. Caruso
Yale University

Gill Sitarenios
MHS Inc.

Manolo Romero Escobar
MHS Inc.

Comments

This is an Open Access article published by Elsevier in Personality and Individual Differences in 2023, available online:

https://dx.doi.org/10.1016/j.paid.2023.112468

Recommended Citation

John D. Mayer, David R. Caruso, Gill Sitarenios, Manolo Romero Escobar, How many emotional intelligence abilities are there? An examination of four measures of emotional intelligence, Personality and Individual Differences, Volume 219, 2024, 112468, https://doi.org/10.1016/j.paid.2023.112468.

Personality and Individual Differences 219 (2024) 112468

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

How many emotional intelligence abilities are there? An examination of four measures of emotional intelligence

John D. Mayer[a,*], David R. Caruso[b], Gill Sitarenios[c,1], Manolo Romero Escobar[c]

[a] Department of Psychology, 15 Academic Way, McConnell Hall, University of New Hampshire, Durham, NH 03824, USA
[b] Office of the Dean, Yale College, 1 Hillhouse Ave, New Haven, CT 06511-8241, USA
[c] MHS Inc., 3770 Victoria Park Avenue, Toronto, ON, M2H 3M6, Canada

ARTICLE INFO

Keywords:
Emotional intelligence
Four-branch model
Emotion perception
Emotion recognition ability
Emotion understanding
Emotion management
MSCEIT

ABSTRACT

The ability model of emotional intelligence (EI) specifies that four related abilities are involved: perceiving emotions, facilitating thought using emotions, understanding emotions, and managing them. Several performance-based assessments have been developed to measure those four abilities. Although some researchers find empirical support for the four abilities, others have argued that emotional intelligence divides into three abilities, two or even a single, unitary ability (Legree et al., 2014; Palmer, Gignac, Manocha, & Stough, 2005). We reanalyzed archival data from four ability tests of emotional intelligence, *Ns* = 503, 5000, 1000, and 2000, conducting item-level exploratory factor models of all four assessments for the first time. Based on those analyses, we suggest possible revisions of the 4-factor model to guide future research and assessment.

Emotional intelligence (EI) has been defined as the "the ability to reason about emotions, and of emotions to enhance thinking" (Mayer, Salovey, & Caruso, 2004, p. 197), and described as consisting of four more specific abilities: accurately (a) perceiving emotions, (b) facilitating thought with emotions (c) understanding emotions and (d) managing emotions (Mayer, Caruso, & Salovey, 2016; Mayer & Salovey, 1997). Empirical evidence indicates that emotional intelligence defined and assessed as a set of abilities is a broad intelligence in the same class as verbal, spatial, and similar intelligences, and that it correlates systematically with such intelligences while exhibiting partial independence from them (Bryan & Mayer, 2021; MacCann, Joseph, Newman, & Roberts, 2014; Schlegel et al., 2019).

Here our focus is on emotional intelligence itself and the subsidiary mental abilities that make it up. Knowledge of specific abilities help determine the test scores that consultants and counselors can responsibly provide to test-takers (Joint_Committee, 2014; e.g., Sinharay, Puhan, & Haberman, 2011). We draw on findings from several key measures of emotional intelligence that have been developed over nearly 25 years to assess the four-area model of emotional intelligence: the Multifactor Emotional Intelligence Scale, (MEIS, Mayer, Caruso, & Salovey, 1999), the Mayer-Caruso-Salovey Emotional Intelligence Test (MSCEIT, Mayer, Salovey, & Caruso, 2002), the Youth Research Version (YRV) of the MSCEIT (MSCEIT-YRV, Mayer, Salovey, & Caruso, 2014) and a forthcoming version (MSCEIT-2; Mayer et al., 2023). All four tests were constructed to assess the four-area ability model of emotional intelligence and each area was assessed by multiple tasks (Mayer et al., 2016; Mayer & Salovey, 1997).

Initial research from the MSCEIT standardization sample supported a 4-factor model corresponding to the four-area model of Perception, Facilitation, Understanding, and Management (Mayer et al., 2002; Mayer, Salovey, Caruso, & Sitarenios, 2003). However, other researchers thought fewer abilities might be involved. Palmer et al. (2005) argued for a 3-factor model that dropped the Facilitation area (see also Fan, Jackson, Yang, Tang, & Zhang, 2010; Mayer et al., 1999). A 2-factor model distinguished between "Experiential" and "Strategic" emotional intelligence (Mayer et al., 2002) combining Perception and Facilitation, on the one hand, versus Understanding and Management. Finally, Legree et al. (2014) argued that just one overall EI ability might exist, and that the remaining factors arose due to method variance from the different response scales the tasks employed: The Perception area of the MSCEIT, for example, employed 5-point Likert scales (e.g., "How much 'happiness' does this face express?"), whereas the Understanding area was multiple choice. Over many studies a consensus seemingly has emerged around the three-factor approach (i.e., dropping "Facilitation";

* Corresponding author.
E-mail addresses: jack.mayer@unh.edu (J.D. Mayer), david.caruso@yale.edu (D.R. Caruso), gsiteranios@yahoo.ca (G. Sitarenios), manolo@itemcinq.ca (M.R. Escobar).

[1] Gill Sitarenios is now at *Acuity Insights;* Manolo Escobar is now a psychometric consultant at *ItemCinq*.

https://doi.org/10.1016/j.paid.2023.112468
Received 27 June 2023; Received in revised form 21 October 2023; Accepted 30 October 2023

Fan et al., 2010; Fiori & Antonakis, 2011; Joseph & Newman, 2010; MacCann et al., 2014; Palmer et al., 2005).

1. Task-level models of emotional intelligence and a proposed item-level alternative

Almost all published research on the factor structure of emotional intelligence has examined the factor structure of the above measures by examining the correlations among their tasks. For example, all the tests had a Perception area that included a "Faces" task, for which participants identified the emotions expressed in a photograph of a face. The Facilitation area consisted of two tasks, one of which, the "Facilitation" task, which asked what mood might help a person plan a birthday party. The Understanding area included a "Blends" task for which people were asked what simple emotions, in combination, are most similar to a more complex feeling such as nostalgia. Finally, the Management area included an Emotion Management task which asked questions such as how best to calm down a friend who just received unsettling news. A full list of the tasks across tests is in Table 1. Collectively, the tasks included in these tests are diverse and wide-ranging, albeit other tasks not included here are also possible, such as an emotion attention regulation factor (Elfenbein and MacCann, 2017).

These task-level models assume that items within a task represent a homogeneous ability. But what if the items were more nuanced than was reflected at the task level? It seems possible that items from different tasks might correlate more closely with one another than items from the same task; if that were the case, models at the item level might reveal a richer picture of EI.

2. Overview of the present studies

In our research here, we analyzed the four assessments, the MEIS, MSCEIT, MSCEIT YRV, and MSCEIT-2, all at the level of their individual items—allowing their items to form factors with one another freely and apart from the tasks to which they had been assigned. Our purpose was forward-looking and focused on the question, "If one were to take a fresh look, what would be the nature and number of abilities one might 'see' in existing tests?"

3. The aim to keep factor models straightforward in the present work

A factor model is simply an approximation of an empirical reality, and more than one model might adequately fit a given dataset. To a degree, factor modeling has become so technical and specialized that it is hard to arrive at a single consensual view of what approach is best (see Markon, 2019; L. K. Muthén and Muthén, 2017; Reise, 2012, for some techniques available). Speaking tongue-in-cheek of these models, E. E. Cureton (1939, p. 287) remarked that a factor analyst, "By the application of higher mathematics to wishful thinking…always proves that his original fixed idea or compulsion was right or necessary…[and] that all other factor-analysts are dangerously insane…".

Our approach here was to keep our analyses straightforward and focused on the essential results by employing a widely-used factor model, the *simple structure* model, with oblique factors. This model allows each test item to load on just one of the factor(s) extracted, and the factors to be correlated. Simple structure models facilitate the interpretation of factors and keep the models consistent across measures. Even so, there were many technical details around the analyses we undertook, and we have reported those in further detail in the accompanying technical supplement (Mayer et al., 2023).

4. Analyses shared in common across the studies

We report four studies, one for each of four assessments in order of publication from 1999 to the present. We kept our plan of analyses as

Table 1
The Emotional Intelligence Tasks Used on the MEIS, MSCEIT, MSCEIT-YRV, and MSCEIT-2.

Task Names: Overall Task	Task Names: Task Division	Descriptions of the Item Forms	Included On this Test or Tests[a]
		Perceiving Emotions	
Music		Identify emotional content of brief musical passage.	MS
Stories		Identify emotion of a character in a story	MS
Pictures	Nature	Identify emotion suggested by a nature photograph	M1
	Designs	Identify emotional content of an abstract design.	MS, M1
Videos		Identify emotion expressed by a person's facial expression from a brief video	M2
Contextual Pics		Identify emotion a person is experiencing from a drawing of the person's posture, and with their face blanked out.	M2
Faces		Identify emotion(s) in a photograph of a person's face.	MS, M1, YRV, M2
		Facilitation of Thought	
Feeling Biases		Imagine a person doing something to you and then how you would evaluate them.	M1
Mood	Energy	Identify energy level of an emotion	M2
Dimensions	Pleasant	Identify pleasantness of an emotion	M2
Changing Contexts/Emotion States		Infer what task a person should work on next to be effective, given a change in their emotions.	M2
Facilitation		Given an emotional state, pick the task at which person will be effective.	M1, M2
Synesthesia/Sensations		Describe an emotion in terms of its physiological sensations	MS, M1, YRV, M2
		Understanding Emotions	
Perspectives		Read a vignette about people and decide how each felt.	MS
Definitions		Identify meaning of emotion words	YRV
Changes/Transitions		Identify how a character's emotion might change after an event.	MS, YRV, M2
Blends		Identify two or more emotions when combined, produce a complex emotion.	MS, M1, YRV, M2
Progressions		Order emotions according to the intensity of their feeling	MS, M1, M2
		Managing Emotions	
Managing Emotions in the Self/ Managing Emotions		Read vignette, imagine yourself in it, choose response most likely to elicit an emotion	MS, M1
Managing Emotions in Others/ Emotion Relations/Emotion Scenarios		Read vignette and choose response most likely to elicit emotion of central character.	MS, M1, M2
Emotion Management, Mixed		Both Self and Other vignettes, mixed	YRV, M2
Picture Panels		View brief visual narrative of people interacting and indicate how to manage a character's emotions.	M2

Notes. a. MS = MEIS, M1 = MSCEIT, YRV = MSCEIT Youth Research Version, M2 = MSCEIT-2.

consistent as possible across studies. We first checked to ensure the relevant datasets and software code generated results that matched earlier published reports. We then tested the fit of a priori confirmatory factor models that have been proposed in the literature to the data. If those failed, we constructed new, tailored factor models to better approximate the obtained data. Details of these models are described in the individual studies; still further information is available in the technical supplement (Mayer et al., 2023, Part 2).

4.1. Scoring approach and conversion to categorical data

The four tests employed different methods of scoring. For the analyses here, we employed the expert-scoring systems for the MEIS and MSCEIT, and veridical scoring for the YRV and MSCEIT-2 because those two methods were most similar across tests. In expert scoring, the proportion of experts who endorse a given alternative as correct is the score for that item alternative. That is, if 0.20 of emotions experts agreed that a “2” on a Likert scale represented the level of happiness in a face, the participant's score for answering “2” was incremented by 0.20. Roberts et al. (2001, p. 202) argued for a further approach they termed “veridical” that would be modeled on certain tasks of standard intelligence tests. To implement their idea, we convened experts to review the relevant literature and create a scoring system to assign, for example, 0 to incorrect, 1 to partially correct, and 2 to fully correct responses. As with other measures of intelligence, items for which the experts could not reach agreement were discarded.

To render data from MSCEIT maximally comparable to the other tests, MSCEIT expert scores, which had fractional values depending upon levels of expert agreement, were converted to categorical data (see Study 2). This resulted in categorical item scores for all four tests (e.g. "0", "1", or "2" as above). This approach reduces artifacts due to the non-normal distributions common to dichotomously (or trichotomously)-scored items (e.g., B. Muthén, 1984; Xia & Yang, 2019).

4.2. Selecting fit criteria across studies

A good factor model should adequately fit empirical findings. In response to a field-wide desire for uniform standards, Hu and Bentler (1999) established criteria for key fit indices: a Root Mean Square Error of Approximation (RMSEA) of <0.06, and Tucker-Lewis (TLI) and Confirmatory Fit Indices (CFI) “near” 0.95. But factor models of test items are far harder to fit than the factor models of test scales or tasks that Hu and Bentler mostly had in mind (e.g., Little, Rhemtulla, Gibson, & Schoemann, 2013). With large numbers of items, the possibility that an item violates a factor model's assumptions are greatly multiplied. To accommodate to this reality, we modified the Hu and Bentler criteria, leaving the RMSEA criterion as is, but with a lowered threshold of the TLI and CFI indices from “near” 0.95 to “at or above” 0.90. This remains a strict criterion for an item-level fit, and, indeed, was challenging to attain here.

In all cases we used the Weighted Least-Squares Mean and Variance Adjusted (WLSMV) method of extraction in Mplus with categorical data specified. For the EFAs we employed a GEOMIN rotation. The key models were single-level (e.g., non-hierarchical) simple structure except where otherwise noted (see details in Mayer et al., 2023).

4.3. Confirmatory tests of a priori models on the full data sets and tailored simple structure models

We first tested four a priori models of EI described earlier (e.g., 1-, 2-, 3-, and 4-factor models), using a confirmatory approach, and then sought to discover whether an alternative model better tailored to the data might fare better. To identify the tailored model, we followed these steps: We first conducted a series of exploratory factor analyses (EFAs) from 1 to 10 factors for each assessment. We used the full sample for each test, rather than reserving a holdout sample for cross-validation because our interest was in consistencies *across* assessments rather than proving a specific model was best for an assessment. Then, we identified “promising” exploratory solutions that both generated interpretable results and exhibited fit indices of TLI and CFI near our 0.900 target (> 0.850 to start). We next reduced the item set, retaining only items that loaded reasonably on (i.e., correlated with) a single factor, and at relatively lower levels on any secondary factor(s). (The simple structure models we employed in both our a priori and tailored CFAs necessitate this constraint to fit). The exact cutpoints for item inclusion varied somewhat across studies owing to the different quality of items across assessments. Finally, we identified a simple structure factor model for each of the assessments that did fit. Simple structure “turns up the contrast” among factors, easing the interpretations and are commonly proposed for that purpose (Thurstone, 1935, 1940), albeit reservations about such models exist (e.g., Ertel, 2011) We describe further specifics of the analyses in the four studies themselves. Our aim was, by the conclusion of the analyses, to identify any consistent patterns across the varied assessments that can inform future research and practice about the more specific reasoning that makes up EI.

5. Study 1: Item-level factor models of the MEIS (1999)

The Multifactor Emotional Intelligence Scale is a 402-item scale that was represented by three or possibly four factors at the task level (Mayer et al., 1999). As with all the assessments here, it was designed to assess the four areas of the ability model of EI. In Study 1, we examine the MEIS at the item level, testing the a priori factor models, as well as a tailored model.

As noted in the “Overview of Studies”, we first tested whether confirmatory a priori 1-, 2-, 3-, or 4-factor models would provide an adequate fit to EI abilities at the item level; and, we also sought to fit an EFA-tailored, simple structure model to identify item groups representing abilities that may have been overlooked in the past.

5.1. Materials and Method

5.1.1. Participants

Participants were 503 adults, as described in the original report, composed of $n = 235$ college students and $n = 268$ part-time college students, corporate business employees, and, career workshop attendees (Mayer et al., 1999, p. 273) divided among $n = 333$ women and 164 men, six unreported.

5.1.2. The MEIS and its tasks

The MEIS is a 12-task scale constructed to assess the ability model's hypothesized four broad areas of emotional intelligence.

The four Perception tasks on the MEIS were Faces, Music, Designs, and Stories (see Table 1 for descriptions of these and other tasks). The two Facilitation tasks were Synesthesia and Feeling Biases. The Understanding area consisted of Blends, Progressions, Transitions, and Perspectives/Relations. Finally, the Managing area consisted of Managing Emotions in the Self and Others. Participants responded to all tasks on 5-point Likert scales (e.g., from emotion absent to present), except for Blends, Progressions, and Transitions, which were multiple choice. Collectively, the tasks contain $K = 402$ items.

5.1.3. MEIS scores

For the analyses here we employed the original expert scoring approach, in which two experts reviewed each item and agreed as to which were correct (assigned 1 point) or incorrect (assigned 0 points).

5.1.4. Procedure

We updated the data file to treat item-level missing data as incorrect *except* if more than half of the items on a task were missing, in which case the item entries were coded as missing, matching the procedure employed for the MSCEIT in 2002.

5.2. Results

5.2.1. Item reduction of the MEIS preliminary to the main analyses

Typical recommendations regarding factor analysis call for a ratio of participants to each indicator variable of 5:1 or 10:1 (Gorsuch, 2015; Nunnally, 1967, p. 436), and take other parameters into consideration as well (e.g., de Winter, Dodou, & Wieringa, 2009; Flora & Flake, 2017). The 1999 analysis of the MEIS employed 503 participants for 12 tasks, a workable 40:1 subjects to task ratio. To model the items, however, would entail a non-viable ratio of 503 subjects to 402 items. We therefore reduced the number of items analyzed by, first, removing 84 items that were overly difficult or easy (correct answers <10 % or > 90 %) because those tend to be less functional and may give rise to difficulty factors (e.g., Aiken, 1979; Schweizer & Troche, 2018). We next counted off the remaining 318 items by threes in order of their position on a given task, yielding three sets, A, B, and C with $K = 117$, 115, and 112 items respectively. All the items from the two shortest tasks (8 items each) were included in all three sets. We then focused on analyses of Set B because EFAs indicated that factor models fit slightly better to that set.

Table 2
MEIS Guided Simple Structure 3-Factor Model, Abridged Solution (Fitting 31 Items of 115 from Item Set B)[a].

Task	Item Description	Factor Loadings		
	Brief Summary of the Item Content	I.	II.	III.
Face	Perceive fear present in fearful face	0.50		
Design	Perceive sadness in a sad design	0.58		
Trans	Perceive fear, as a plausible reaction	0.38		
Face	Perceive disgust in disgusted face	0.48		
Music	Perceive surprise in happy, lively music	0.44		
Persp.	Identify lack of embarrassment in a parent's reaction to child	−0.30		
Design	Identify lack of disgust in a neutral design	−0.45		
Tran	Identify lack of surprise after an emotion transition	−0.43		
Design	Identify lack of fear in response to a warm design	−0.44		
Face	Identify lack of fear in non-fearful face	−0.44		
Music	Identify lack of sadness in percussive music	−0.47		
Bias	Perceiving a coworker as calm after they praised you		0.56	
Design	Sadness in a design of two abstract human shapes		0.46	
Design	Sadness in response to purple-grey design		0.53	
Manag. Other	Help a friend feel better by reconnecting with significant. other		0.39	
Design	Sadness in an oily-black metal-like design		0.45	
Bias	Attributing hastiness to relative who made you feel threatened		0.48	
Design	Surprise in prismatically colorful pattern		0.41	
Music	Happiness in sad, contemplative music		0.42	
Synest.	Associating "high" (v. "low") to contented satisfaction		0.34	
Story	Feeling guilty after frustrations with ex-spouse		0.30	
Synest.	Associating "dark" with embarrassment		0.34	
Persp.	Embarrassment at parent's child's bullying behavior			0.49
Prog.	Feeling rage at point of feeling angrier and fearing losing control			0.55
Blend	Remorse as combination of guilt, regret, and sadness			0.51
Persp.	Feel relief after telling friend you were promoted but not the friend			0.43
Prog.	Excitement at the point of feeling livelier and livelier			0.51
Blend	Optimism as combining pleasure and anticipation			0.33
Persp.	Dog owner-guardian angry at themselves for carelessness			0.38

Notes: a. Items are in order of presentation in their task. Only the three top-loading items for a task on a specific factor are included in this abridged table; 29 items of 31 items appear. See text for details as to Set B.

5.2.2. Fits of a satisfactory model to the $K = 115$ items of Set B of the MEIS

We first tested confirmatory, a priori 1-, 2-, 3- and 4-factor simple structure models to the Set B items. All failed. The four-factor model, which fit best, had CFI and TLI of 0.610 and 0.602. Details are in Mayer et al. (2023).

We next conducted exploratory factor models of the MEIS, focusing on 3, 4, and 5 factor models. Although the fit statistics were reasonable for those models (e.g., for 4 factors, χ^2 was 6339.36 (df = 6101), RMSEA of 0.009 and CFI and TLI of 0.928 and 0.921), many items loaded on multiple factors—which rendered them problematic for fitting a simple structure—or failed to load at all. After experimenting, we concluded the items required a criterion for retention of $r > |0.35|$ on the primary factor, and $r < |0.20|$ on any secondary factor, to arrive at a fitted solution. Relaxing those criteria (i.e., primary $r = |0.25|$) led to much poorer fits. Applying the criteria there were no possible solutions beyond a 4-factor solution, and the four-factor model was non-viable in that it yielded a non-positive definite matrix. This left 34 items.

The fit of the 3-factor solution to the 31 remaining items of the original 115 (3 items were dropped based on modification indices) met our criteria with an RMSEA of 0.016 and CFI and TLI of 0.913 and 0.906. Table 2 contains a slightly abridged version of this solution.

We labeled Factor 1 *Endorsement Bias*, because it reflected some participants' tendency to endorse the presence of emotions in a stimulus (e.g., a face or a musical piece) whereas other participants endorsed their absence. We labeled Factor 2 *Connecting Emotional Features* because it represented the ability to perceive components of emotion and to connect them across modalities, as in connecting emotion elements to a person's behavior (e.g., Bias), or inferring them from an abstract design (e.g., Designs), or generalizing them to other sensory modalities (e.g., Synesthesia). Factor 3, Understanding, exclusively loaded items of the Understanding area.

Further supporting these interpretations, Endorsement Bias correlated just $r = 0.12$ with the overall MEIS, whereas the Understanding factor correlated $r = 0.66$ with the Understanding area, and Connecting Emotions correlated substantially with the Emotional Perception and Facilitation areas, $rs = 0.57$ and 0.43. The sum of the two substantive factors (two and three) recovered the full 402-item MEIS test scores at $r = 0.69$ with just the 29 items shown in the table.

5.2.3. A check of alternative models

Given that Item Set B contained just one subset of MEIS items, we also examined the EFAs for Sets A and C. Their 3-factor EFAs both contained "Emotion Present versus Absent" and "Understanding" factors. For Set A, the remaining factor represented "Emotion Management" or possibly a more general "Interpersonal Emotion Understanding", as it included not only emotion management items, but also Perspectives task items, which involved understanding others' emotions. For Set C, the remaining factor represented Emotion Perception accuracy.

5.3. Discussion of Study 1

Our analyses of the MEIS began with item reduction to create three smaller sets of items. We focused on "Set B" which seemed most amenable to good fits; nonetheless, fitting a tailored solution to the data set left just 31 items because so many items were non-specific in their loadings. Set B (and Sets A and C) all contained both an Understanding factor and a method factor reflecting participants' tendencies to endorse emotions as present or absent in pictures. Set B also contained a new factor we called Connecting Emotional Features that reflected the capacity to connect emotional qualities across modalities (e.g., in abstract designs and synesthetic experiences). A simple sum of the 18 highest-loading items on the Understanding and Connecting factors correlated with overall MEIS scores (based on 402 items) at $r = 0.69$, suggesting the centrality of these factors to performance on the overall test.

6. Study 2. Item-level factor models of the MSCEIT (2002)

Next analyzed was the MSCEIT, a 141-item scale of which 122 items are scored. Researchers who analyzed the test at the task level argued variously for 4-, 3-, 2- and 1-factor solutions (e.g., Fan et al., 2010; Palmer et al., 2005). In Study 2, we examined the test at the item level. We again began by testing a priori confirmatory 1-, 2-, 3-, or 4-factor models at the item level; and we also sought to fit an EFA-tailored, simple structure model to see if the items might reflect abilities that had been overlooked in the past.

6.1. Material and Method

6.1.1. Participants

Participants were 5000 adults who constituted the standardization sample of the MSCEIT. The exact composition of the sample is described in the test manual, but, in short, matched a North American census at that time, and represented diverse educational, ethnic, and (adult) age groups, split between women and men (Mayer et al., 2002).

6.1.2. The MSCEIT and its tasks

The eight tasks of the MSCEIT were, for Perceiving Emotions: Faces and Pictures; for the Facilitating Thought area: Facilitation and Sensations; for the Understanding area: Changes and Blends; for Management: Emotion Management (i.e., Self-Management) and Emotional Relations (i.e., Managing Others). All tasks employed Likert scales, excepting those in the Understanding area, which were multiple choice. The tasks are described in Table 1.

6.1.3. MSCEIT scores

As noted in the "Overview of Studies," we employed expert-scoring of the individual items, converted to an ordered categorical form. The original MSCEIT expert scoring was based on 21 international emotion experts who each identified what they considered to be the best answer to each item. Their proportion of endorsements were used to assign corresponding credit to participants. So, if the experts split 60/40 between two reasonable alternatives, participants who endorsed the first would receive a score of 0.60; those participants who endorsed the second, a score of 0.40 and any other alternative would be scored 0 (Mayer et al., 2003). The proportional item scores of the MSCEIT were converted to categorical form to render its scores commensurate with the other tests. For the aforementioned expert-scored example, for instance, 0-credit answers were converted to 1, 0.40 answers to 2, and a 0.60 answer to 3. Owing to the variation across item ratings, some items took on just two such values (1,2), others as many as 6 values (1 through 6), where 1 was no credit/omitted.

6.1.4. Procedure

The 5000-person standardization sample was archival and had been stored as an Excel file since its original collection by the publisher in 2001.

6.2. Results

6.2.1. Fits of the a priori factor models to the MSCEIT at the item level

Tests of the a priori confirmatory factor models of the 1-, 2-, 3- and 4-factor simple structure models all failed at the item level, with the fit for the highest-performing four factor model at $\chi2(7253) = 44{,}812.14$, and RMSEA, CFI and TLI of 0.032, 0.761, and 0.756.

6.2.2. The tailored model to fit to the MSCEIT

An initial exploratory factor analysis of the 122 scored items indicated a window of possible models that extended from 4- to 7-factor models. Examining those possible models turned up some surprises: By just 4 factors, the Perception area Faces and Pictures tasks split into two factors, as had the first and second halves of the Changes items (which were phrased differently). These splits remained through the 7-factor solution and it took until the 7-factor solution for Facilitation to separate from the Emotion Management items. We therefore chose to model the MSCEIT based on the 7-factor model, which fit reasonably well to start ($\chi2(6548) = 17{,}618.01$, RMSEA, 0.02, and CFI and TLI of 0.929 and 0.920) (see Mayer et al., 2023, for details).

6.2.3. Fitting the simple structure 7-factor model

We set criteria for item retention based on the 7-factor solution at $r > |0.30|$ on each item's primary factor and selected only those items for which their secondary loading was "appreciably lower" (e.g., an $r > |0.10|$ difference), reducing the items to 94. From there, we dropped a further 6 items that violated the simple structure model, based on modification indices. The final fit was $\chi2(3719) = 15{,}160.05$; RMSEA, 0.025, and CFI and TLI of 0.904 and 0.901.

We engaged in a second attempt to fit the four-factor a priori model, this time to the same 88 items in tailored model (rather than the initial 122); it failed again at ($\chi2(4271) = 33{,}151.39$, and RMSEA, CFI and TLI of 0.037, 0.776, and 0.770.

6.2.4. Interpretation of the 7-Factor EFA-guided constrained factor analytic model

As indicated in Table 3, Factor 1, Faces, predominantly loaded items from the Faces task. Factor 2, Nature Pictures, loaded the nature photographs of the Pictures task. Factor 3, Connecting Emotion Features, loaded items from the abstract designs portion of Pictures and Synesthesia. Factors 3, 4 and 6 each represented an Understanding Task. It was surprising and unclear why they did not merge (adding to the mystery, two employed the same multiple choice response scales). Factor 5, Management, represented both Managing Emotions tasks (Emotion Management and Emotional Relations). Finally, factor 7, Facilitation, loaded items from the Facilitation task. The independence of Facilitation was relatively unsurprising given the uniqueness of its items relative to Synesthesia (with which it was paired). Despite the division into 7 factors, all exhibited moderate correlations among themselves, typically between $r = 0.25$ to 0.60. And, although our a priori 4-factor model had failed on the 88 items, a further *hierarchical* 4-factor model, which nested the 7 factors and 88 items in the four areas almost fit, indicating some continuity with a 4-factor approach ($\chi2(3728) = 15{,}711.99$, RMSEA of 0.025, and CFI and TLI of 0.899 and 0.897, with a mild Heywood case). Note, however, that the model depended on rearranging the items according to the present 7-factor model. Finally, the composite score of the 7 factor-based scales employing the 37 (of 88) items of Table 3 recovered the full MSCEIT test score at $r = 0.89$, indicating they captured the gist of the overall test.

6.3. Discussion of Study 2

Rather than the foundational four-factor (or three-factor) solutions found by researchers modeling the MSCEIT *tasks*, the item level analyses here suggested a 7-factor solution to meet our criteria of a reasonable fit (this included 88 of the test's 122 scored items). The solution included *two* factors for emotional perception: one for Faces, and the other for the Nature Pictures of the Pictures task. A Connecting Emotional Features factor again emerged that combined Synesthesia with Abstract Designs items (from the Pictures task) and Understanding and Management factors emerged. The last two factors represented a second subset of Changes items, and all the Facilitation items. The fact that the Facilitation task (and the second set of Changes items) split off through most analyses was unexpected and will be further considered in the General Discussion.

Table 3
MSCEIT Guided Constrained Factor Analysis for a 7-Factor Model, Abridged Solution[a].

Item	Description	I	II	III	IV	V	VI	VII
Faces	Surprise in a neutral face	0.72						
Faces	Excitement in somewhat sad face	0.74						
Faces	Excitement in tired, sad face	0.71						
Faces	Fear in happy face	0.69						
Faces	Disgust in happy face	0.73						
Pictures	Sadness in pretty water scene		0.59					
Pictures	Fear in pretty water scene		0.59					
Pictures	Disgust in pretty water scene		0.85					
Pictures	Anger in picture of a desolate bush		0.65					
Pictures	Disgust in picture of a desolate bush		0.60					
Pictures	Sadness in lively cartoon-like design			0.86				
Pictures	Fear in lively cartoon-like design			0.77				
Pictures	Disgust in lively cartoon-like design			0.81				
Sensations	Surprise as “cold, slow, and sharp”			0.61				
Sensations	Afraid as “loud, large, delicate, and dark green”			0.55				
Sensations	Calm in “closed, dark, and numb"			0.51				
Changes	Person's partner seemed perfect, adored him				0.52			
Blends	Concern blends anxiety, caring, etc.				0.50			
Blends	Anticipating pleas. Reflects optimism				0.59			
Blends	Shame and embarrassment form humiliation				0.51			
Blends	Sadness and satisfaction as nostalgia				0.59			
Emot. Mng.	Preserve good mood by thinking positively					0.55		
Emot. Mng.	Preserve good mood, reject incorrect choice					0.66		
Emot. Mng.	Guide anger well, rejecting inappropriate action.					0.61		
Emot Rel.	Help departing friend to sustain relationship					0.55		
Emot Rel.	Fail to help departing friend, showed disapproval					0.64		
Emot Rel.	Parents' counterproductive approach to school					0.53		
Changes	Secure to depressed due to illness						0.71	
Changes	Displeased to resentful due to unfair treatment						0.84	
Changes	Anger to guilt owing to inappropriate behavior						0.87	
Changes	Liking to despising due to betrayal						0.73	
Changes	Love to security due reciprocal feelings						0.97	
Facil. Thought	Plan party with joy							0.73
Facil. Thought	Reject boredom as helping plan party							0.53
Facil. Thought	Frustration as unhelpful to music composition							0.50
Facil. Thought	Reject sorrow as helping epicurean pursuits							0.75
Facil. Thought	Correctly reject anger to establish transition plan							0.51

a. The factor model was fit to 88 items of the original 122 MSCEIT items scored. The table shows 37 high loading items in order of their presentation in their task. The table shows 5 to 6 items per factor. When a factor loaded more than one task, items with loadings of 0.3 or above were included up to three items for each task.

7. Study 3. Item-Level Factor Models of the MSCEIT-YRV (2014)

The MSCEIT-YRV is a 6-task, 97-item assessment designed to measure the four areas of EI tailored to assessing 10 to 18-year-olds.

7.1. Materials and Method

7.1.1. Participants

The archival, normative sample examined here was composed of 1000 youth distributed among 3 age groups of 10–12, 13–15, and 16–18 years of age, with each age bracket evenly split between male and female participants. They represented an ethnically diverse group including White, Asian, Black, Native American, and multiethnic participants, also with parents of diverse socioeconomic status. Details can be found in the test manual (Mayer et al., 2014).

7.1.2. The MSCEIT-YRV and its tasks

The MSCEIT-YRV was designed to be shorter than earlier tests so as to accommodate younger test takers (Mayer et al., 2014). The Perception area was measured by Faces of young people; the Facilitation area was assessed by Synesthesia; the Understanding Emotions area by Emotion Definitions, Transitions, and Blends; Management, by a combination of items addressing self- and other- emotion management. All tasks utilized Likert scales except the Understanding tasks, which were multiple choice.

7.1.3. MSCEIT-YRV scores

The MSCEIT-YRV employed veridical scoring. For veridical scoring, panels of emotions experts were convened who, after studying a scoring booklet summarizing relevant emotion research, were to reach a consensus as to the correct answer or answers to each test item. An item was typically scored “0” for incorrect, “1” for partial credit, and “2” for credit, although for some items, intermediate scores were permitted (e.g., “1.5”). If the experts could not agree on a score, the item was dropped.

7.1.4. Procedure

The 1000-person standardization sample was archival and had been stored as an Excel file since its original collection by the publisher in 2014. Because the veridical scoring sometimes included intermediate, fractional values, the data were converted to ordered categorical form (whole numbers) for the analyses. In the MSCEIT-YRV original scoring, a given Face's several Likert scales (one for each emotion that might be expressed) were combined into a discrepancy score to index how close the respondent's answers were to the ranking of Likert scales by the experts. This approach had been implemented to mitigate concern over response biases present on the perception tasks of the MEIS and MSCEIT. To adhere to our process here, we analyzed each response scale (i.e., item) individually.

7.2. Results

7.2.1. Fits of the a priori factor confirmatory models to the MSCEIT-YRV at the item level

Our confirmatory tests (CFAs) of the a priori 1- through 4-factor models at the item level all failed. The 3-factor model, which dropped the 24 Facilitation items on the scale, fit best at χ2(4553) = 5373, and RMSEA, CFI and TLI of 0.033, 0.806, and 0.800.

7.2.2. Creating the tailored model

An exploratory factor analysis of 97 MSCEIT-YRV items suggested a window of possible models that fit somewhere between four and six factors. That said, the 4-factor model (and lower dimensions) omitted a clean Perception factor, which did not emerge until 5 factors. We therefore selected the 5-factor version.

To create the simple structure model, we selected 68 items from the original 97 that loaded $> |0.25|$ on their primary factor and about $r = 0.1$ higher than on any secondary factor. Dropping an additional three items that violated the overall model led to a model fit with a χ2(2005) = 3527.05, and RMSEA, CFI and TLI of 0.028, 0.907, and 0.903. An abridged version of the model is indicated in Table 4. We also refit the a priori models after our tailored item reduction and the 3-factor model met the criteria for a fit, after removing all the Facilitation area items (i.e., the Synesthesia task items), with χ2(1172) = 2152.62, and RMSEA, CFI and TLI of 0.029, 0.919, and 0.915). The next best fits of the a priori models were not close.

7.2.3. Interpretation of the 5-Factor EFA-guided constrained factor analytic model

Factor 1, Perception, loaded the Faces task; Factor 2, Synesthesia, loaded many of the Synesthesia items. Factor 3, Understanding, included items from all three Understanding tasks. Factor 4, Management, loaded management-task items exclusively. The interpretation of Factor 5, Attention, which combined items from Synesthesia and Faces was murkier: It involved the correct dismissal of very unlikely answers on the two tasks. We regard it as reflecting an attention factor, although it is possible it reflects Connecting Emotions. The five factors exhibited a wide range of positive correlations (with one near-zero exception), from $r = 0.17$ to 0.82. A sum of the factor-based scales composed just of the 29 items in Table 4 recovered the full MSCEIT-YRV scores $r = 0.90$. The technical supplement has further details (Mayer et al., 2023).

7.3. Discussion of study 3

The 5-factor model of the YRV recovered factors representing Perception (Faces only), Understanding, and Management, as well as a factor exclusively consisting of Synesthesia items that might be either Facilitation or the newly proposed Connecting Emotions factor. An additional factor may represent “attention” as it consisted of very easy items; because it combined some Face and Synesthesia items it might be another candidate for the Connecting Emotions factor. We hoped for further information from analyses of the MSCEIT-2

8. Study 4. Item-level factor models of the MSCEIT-2 (2023)

The MSCEIT-2 is an 83-item scale that, unlike the prior tests, was developed at the item level. It contains 12 tasks, distributed among the four theoretical EI areas of Perceiving, Facilitating, Understanding, and Managing.

8.1. Materials and Method

8.1.1. Participants

The 3000-person standardization sample for the MSCEIT-2 was drawn to be representative of North American census data and was collected by MHS of Toronto, Canada.

8.1.2. The MSCEIT-2, its scales, and its scoring

On the MSCEIT-2, the Perceiving area is measured with Contextual Pictures, Faces, and Videos; Facilitating Thought is assessed with Changing Contexts, Facilitation, Emotion Dimensions, and Sensations (Synesthesia); Understanding with Blends, Changes, and Progressions; and Management with Picture Panels and Emotion Scenarios (see Table 1 for details as to tasks). Items were multiple choice in format; the test employed veridical scoring with “0” no credit, “1” partial credit, and “2” full credit.

Table 4
MSCEIT-YRV Guided Constrained (Confirmatory) Factor Analytic 5-Factor Model, Abridged Solution[a].

Task	Item Description	1	2	3	4	5
Face	Surprise in happy face	0.63				
Face	Disgust in mildly sad face	0.70				
Face	Surprise in excited, happy face	0.60				
Face	Disgust in neutral, calm face	0.69				
Face	Surprise in happy face	0.63				
Synest.	Hot as “angry”		0.76			
Synest.	Red as “angry”		0.76			
Synest.	Slow as “sad”		0.72			
Synest.	Heavy as “sad”		0.83			
Synest.	Sadness as “cold, slow, heavy”		0.76			
Defn.	Unintentional hurt leads to guilt			0.58		
Defn.	If someone has a desired possession you feel envy			0.67		
Defn.	Self-approval describes self-satisfaction			0.67		
Chngs.	Hopeful upon seeing possible romantic interest			0.83		
Chngs.	Worried about being caught			0.84		
Chngs.	Admired by those who liked them			0.88		
Blends	Aggressiveness blends anger and anticipation			0.40		
Blends	Remorse includes guilt and regret			0.40		
Blends	Calmness includes relaxed and secure			0.72		
Mang.	Use distraction and lose focus on goal				0.72	
Mang.	Use alternate distraction and lose focus on goal				0.80	
Mang.	Become preoccupied before congratulating another				0.94	
Mang.	Manage worries by dwelling on them				0.74	
Mang.	Motivate team through quiet contemplation				0.60	
Face	Happiness in happy face					0.41
Face	Disgust in happy excited face					0.41
Synest.	Sadness as “warm and light”					0.83
Synest.	Worry as “warm and light”					0.87
Synest.	Guilt as “warm and light”					0.84

a. The factor model was fit to 65 items of the original 97 MSCEIT-YRV items. The table shows 29 of the 65 items fit in order of their presentation in their task, with up to 3 highest items above 0.3 selected from each of the tasks that loaded on a factor.

8.1.3. Procedure

The MSCEIT-2 data was collected by a subcontractor with data collection continuing until all demographic groups were filled to a criterion that matched a North American census.

8.2. Results

8.2.1. Fitting factor models to the MSCEIT-2

For the first time, the a priori confirmatory factor models of the test fit. A 1-factor model, dropping 4 items from Faces based on modification indices, met criteria with $\chi2(3002) = 6559.66$, RMSEA, 0.020, and CFI and TLI of 0.924, and 0.922. The fits of the factor models improved from there. The 2-factor solution was at criterion (except the TLI dropped to 0.897; we did not explore item drops), and the 3- and 4-factors readily fit the data with no item drops. The 4-factor a priori model fit with $\chi2(3314) = 6788.99$, and RMSEA of 0.019, and CFI and TLI of 0.933 and 0.931. That said, EFA-guided 3- and 4-factor simple structure factor models fit slightly better, with the 4-factor version, for example, at $\chi2(3314) = 6204.60$, the RMSEA at 0.017, and CFI and TLI of 0.944 and 0.943. Both the a priori 4-factor model, and the 3- and 4-factor guided models are reported in abridged form in Table 5. The specific items of the table were selected on the basis of the a priori 4-factor model, and we carried the items across the table to indicate their loadings on the 3- and 4-factor guided models.

8.2.2. Interpretation of the 4-factor and 3-factor models of the MSCEIT-2

The interpretation of the factor models of the MSCEIT-2 are relatively self-evident. The four factors of the 4-factor a priori model represent, in turn, Perception, Connection, Understanding, and Management. On the 3- and 4-factor guided models, Contextual Pictures moves out from Perception to either Understanding in the 3-factor or Connecting Emotions in the 4-factor version. Changes seems to belong equally to Understanding and Management. One hypothesis is that understanding the changes that bring about an emotion could be helpful in managing emotions. The factors of all three solutions were moderately to highly correlated with one another in the model. The 28 of 83 items shown here for the a priori solution recovered the full test score $r = 0.92$ in this data set.

8.3. Discussion of Study 4

For the first time among these tests, the *a priori* models fit the MSCEIT-2 data at the item level, likely because the test was developed at the item level. Nonetheless, guided models fit slightly better. We discuss these and other matters in the general discussion.

9. General discussion

Our aim in this paper was to take a new look at the structure of emotional intelligence abilities by delving into the individual items of four tests of varied tasks, across multiple samples. All tests had been constructed to measure accuracy in four areas of emotional intelligence: (a) emotional perception, (b) facilitating thought, (c) understanding emotions, and (d) managing emotions. Each test employed between six and twelve tasks, with each task measuring a specific area.

9.1. Examining factors across samples and across tests

One strength of the analyses here was that we examined emotional intelligence not only across different samples of test-takers, but also across different samples of test items. One consequence was that we saw considerable variability across the tailored factor models we employed to fit EI across measures. It took from 3 factors (for the MEIS) to 7 for the MSCEIT to reach good fits to the data. Over the near 25-year-period across which the tests were developed, item analyses of the scales became more common and, consequently, our models were able to fit a

Table 5
MSCEIT-2, 4-Factor A Priori and 3-Factor Guided Factor Models, Abridged Solutions.

	4-Factor A Priori Solution				3-Factor Guided Version			4-Factor Guided Version			
Task	I	II	III	IV	I	II/III	IV	I	II	III	IV
Contextual Pictures	0.36					0.34			0.36		
Contextual Pictures	0.34					0.31			0.33		
Faces	0.81				0.83			0.83			
Faces	0.95				0.98			0.98			
Faces	0.68				0.70			0.70			
Videos	0.32				0.32			0.32			
Videos	0.35				0.35			0.34			
Changing Contexts		0.34				0.32			0.34		
Changing Contexts		0.43				0.40			0.43		
Sensations		0.32				0.30			0.32		
Emotion Dimensions		0.61				0.57			0.61		
Emotion Dimensions		0.68				0.63			0.67		
Emotion Dimensions		0.54				0.50			0.54		
Blends			0.45			0.46				0.45	
Blends			0.42			0.43				0.42	
Blends			0.52			0.54				0.52	
Changes			0.59				0.60			0.59	
Changes			0.66				0.68			0.66	
Changes			0.61				0.62			0.61	
Progress-Excitement			0.57			0.58				0.57	
Progress-Excitement			0.58			0.60				0.58	
Progress-Pleasantness			0.70			0.72				0.70	
Manage-Scenarios				0.67			0.67				0.67
Manage-Scenarios				0.76			0.75				0.76
Manage-Scenarios				0.61			0.60				0.61
Manage-Text				0.50			0.50				0.50
Manage-Text				0.45			0.44				0.45
Manage-Text				0.70			0.70				0.70

a. The table shows 28 high loading items of the 83 items for the 4-factor a priori model with up to 3 items loading above 0.3 selected from each of the tasks that loaded on a factor. The loadings for the same items (i.e., selected on the basis of the a priori model) also are indicated for the tailored 3- and 4-factor models (fitting the same 83 items). Items are in order of presentation in their task.

progressively larger proportion of the items of a test, from just over a quarter of the Set B items of the MEIS to all items on the MSCEIT-2.

We also note that even slight modifications in item format could result in a task splitting across two factors, as occurred on the Changes and Pictures tasks of the MSCEIT at the item level. Additional method variance arose, no doubt, from whether the items employed predominantly Likert scales, as had the MEIS, or consisted of multiple choice response formats, as with the MSCEIT-2 (see also, Legree et al., 2014).

One positive aspect of this variability across samples and tests was that whatever factors showed up consistently across scales would likely be key to representing emotional intelligence. One limitation of the current findings was that the four measures and their diverse tasks do not include all possible EI tasks such as "Emotional attention regulation" and "memory for emotional contents" (Elfenbein and MacCann, 2017), which might potentially yield further factors. Inclusion of such tasks may well yield a further expansion regarding "how many abilities" there are in ability measures of EI.

9.2. A revised view of emotional intelligence abilities

The findings here across samples and items, as we interpret them, are summarized in Fig. 1. Fig. 1 divides emotional intelligence (top) into five broad factor-based and conceptual areas (second row), from Perceive to "Not Yet Classified." Beneath each of the five areas are the specific tasks that may belong to it. We describe each of these five areas below.

9.2.1. Accuracy in emotion perception: one ability or many?

The dotted line circling the Emotion Perception area ("Perceive") indicates a finding across our analyses at the item level that Perception items often split into more than one factor. Putting aside the obvious method factor on the MEIS, this was clear on the original MSCEIT and on the MSCEIT-2 as well. From Buck's (1984, p. 281), observation that different measures of emotion recognition accuracy were largely uncorrelated, through the work of Schlegel, Grandjean, and Scherer (Schlegel, Grandjean, & Scherer, 2012), to the present (Schlegel et al., 2019) several accounts of the possible divergences have been proposed, but perhaps the modality, e.g., still faces versus video faces versus voices versus nature pictures, is key, after method factors are mitigated (e.g., Lewis, Lefevre, & Young, 2016; Schlegel, Boone, & Hall, 2017, Table 5).

On the MSCEIT, perception split into (a) Perception in Faces (b) Perception in Nature, and (c) a third Perception area of Abstract Designs, which may better be regarded as part of the revised "Connecting Emotion Features" factor discussed below. On the MSCEIT-2 the Perception factor loaded Faces items predominantly. Although the Videos (of Faces) and Contextual Pictures items also loaded on the same factor, the loadings were so low that it was not clear they really belonged and, indeed, alternative 3- and 4-factor solutions placed Contextual Pictures on the Understanding factor (3-factor solution) and Connecting Emotions (4-factor solution). The findings are consistent in suggesting that directly perceiving emotional expressions in the still faces is distinct to a degree from perceiving emotions in artwork and music, for which some connections may need to be established between colors, sounds, and textures in the artistic products, on the one hand, and emotions, on the other—as we discuss next. But even here there is fragmentation commonly found among direct emotion perception tasks across different modalities such as still faces versus videos.

9.2.2. Observations regarding the ability at connecting emotion features

The second factorial area in Fig. 1 is "Connecting Emotions" ("Connect" in Fig. 1); it is encircled by a dashed line indicating our proposal regarding its existence is tentative as it has not been demarcated quite this way before. Connecting Emotions seems to involve detecting

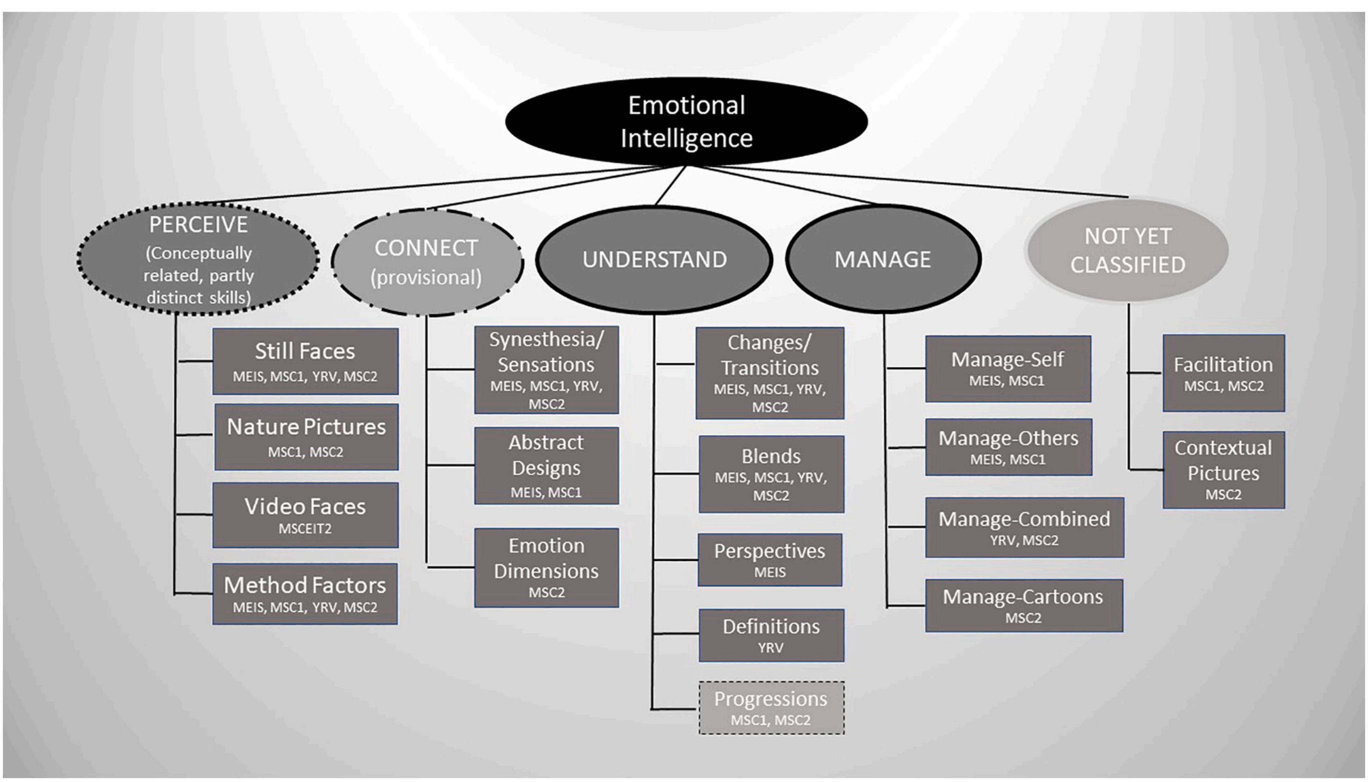

Fig. 1. A revised conception of the abilities of emotional intelligence. In this reconfigured view of emotional intelligence, we infer the possibility of four groups of abilities and one group of "unclassified", with indicator tasks beneath them. Understanding and Management are the most reliable factors. Perception seems to consist of multiple factors and Connecting Emotions is a provisional group. Beneath each group are the tasks that describe it, followed by the test on which it appeared, i.e., the 'MEIS', the Multifactor Emotional Intelligence Test; 'MSC1', the original MSCEIT; the 'YRV', MSCEIT-Youth Research Version, and 'MSC2', the MSCEIT-2.

emotional features and then connecting them to other sensory phenomena (e.g., through synesthetic connections). The factor appeared on both the MEIS and original MSCEIT, where it consisted of items from emotion perception in abstract designs and items from Synesthesia. The same factor arguably appeared on the MSCEIT-2, where it combined Synesthesia and Emotion Dimensions. It appeared more robust than the original Facilitation factor and drew on some of the tasks from that area, but less on the items from the Facilitation task itself which, as we examine below, formed its own factor.

9.2.3. Factors of emotional understanding and emotional management

Given the remarks concerning Emotion Perception and Connections, we note that the next two areas in Fig. 1, Understanding and Management, are fully consistent with theory in the area. An Emotion Understanding factor appeared across all four studies, albeit several Understanding tasks split away from the main factor on the MSCEIT. An Emotion Management factor appeared in analyses of the MEIS (Set A items), MSCEIT, YRV, and the MSCEIT-2.

9.2.4. Not-yet-classified items

There was less consistent evidence of a Facilitation factor across these studies. As noted earlier, this theoretically-conceived of factor frequently failed to find empirical support (e.g., Fan et al., 2010; Palmer et al., 2005), although analysis of item parcels provided some support for the factor (Mayer et al., 2002). In the present work a pure Facilitation factor was not obtained and the Facilitation task formed its own factor on the original MSCEIT, distinct from the others. For that reason, we have placed the task under "Not-Yet-Classified" areas.

9.2.5. Method factors

The MEIS and MSCEIT, and possibly YRV (Factor 5) exhibited method factors as well. The Faces task in particular is readily influenced by response biases of various sorts; on the MEIS, the Faces task measured the tendency to endorse "emotion present" versus "emotion absent" alternatives rather than true ability. And the YRV exhibited a factor representing "Rejecting Implausible Alternatives," which might also arguably reflect attention or a method factor. These method factors were placed under the "Perceive" group.

9.3. The divergence between the task and item level: which is best?

Considerable numbers of analyses indicate that 1-, 2-, 3-, and 4-factor models can successfully model these tests at the task level. But with the exception of the MSCEIT-2, those models failed to fit the tests at the item level. Two explanations seem possible. The first concerns the mathematical models themselves: Perhaps the fit statistics for item level analyses are themselves too restrictive and unbending to accommodate item-level performance and therefore the fit criteria ought to be greatly relaxed in such cases. The concern with allowing a looser fit, however, would be to make it harder to distinguish good from bad models. (One also could argue against constrained, simple structure models in this regard: Perhaps allowing items to load on multiple factors is best).

An alternative possibility is that the mathematical models performed more-or-less adequately, and that tests simply need to be designed and evaluated at the item level from the start. After all, the human part of the equation—the test taker—encounters tests on an item-by-item basis, albeit they also work through materials task-by-task. It may be that intelligence tests would be more robust if they were more often developed at the item level, as was the MSCEIT-2.

9.4. Relation to prior analyses of the factor structure of emotional intelligence

Our depiction of the relations of emotional intelligence abilities are, in certain respects, supportive of research by Palmer, Gignac, and Fan, who had argued for a three-factor representation (Fan et al., 2010; Palmer et al., 2005). All three areas of their three-factor solution were found here at the item level: Perception, Understanding, and Management. That said, the item-level findings here indicated that Perception may itself divide into distinct factors of still-face perception, video-face perception, nature perception and possibly other factors. And the 3-factor solution may overlook both the Connecting Emotion factor proposed here, and the as-of-yet unplaced tasks such as Facilitation. As Legree et al. (2014) argued it might, a single factor solution did fit the newest test reasonably well, perhaps because the MSCEIT-2 had been developed at the item level and employed multiple-choice response scales throughout. That said, higher dimensional solutions fit somewhat better.

9.5. Strengths and limitations

The analyses here were conducted on a series of distinct tests developed over two decades, each on its own sample ranging from $N =$ 503 to 5000. One limit to this diversity of items and samples was that the scales were based on the same four-ability model. There are other ability scales of emotional intelligence such as the Geneva Test of Emotional Intelligence with other items that might also be examined in this way (Schlegel & Mortillaro, 2019). The "Voices" task on the Geneva test might represent another instance of what appears to be a diverse multifactorial realm of emotion perception. In addition, after testing the a priori factor models of each test at the item level, we mostly focused on constrained, simple-structure factor models of the data, with the exception of a hierarchical model of the MSCEIT. Other approaches might well have yielded somewhat different findings, including approaches such as network analyses (McGrew, Schneider, Decker, & Bulut, 2023).

10. Conclusions and implications for assessment

Theory is a powerful tool in the sciences but it is also important to adjust theories as new empirical findings improve our understanding. The theory of emotional intelligence has been associated with two key ideas: That emotional intelligence can be assessed as a mental ability using intelligence tests, and further, that the intelligence divides into four more specific abilities: accuracy at perceiving emotion, facilitating thought, understanding feelings, and emotion management. Of the first claim—that EI is an intelligence and can be assessed with intelligence tests—there is little doubt: Research from the MEIS forward demonstrates that the four-area model of emotional intelligence can be operationalized as a series of ability-based assessments (MacCann et al., 2014).

The research here, however, indicates that the four-area model could benefit from some reconsideration. The findings indicate the robustness of scores reflecting Understanding and Managing abilities, but that the Emotion Perception area may be more multi-dimensional than the original theory indicated, and that a "Connecting Emotional Features" score may better represent certain tasks than the earlier-conceived "Facilitating Thought" area. The findings make room, as well, for as-of-yet unclassified tasks (see Fig. 1). Collectively, the results encourage us to rethink EI in ways that have plain implications for future measures.

To the degree such new understanding closer approximates the reality of mental abilities in this area, it can inform applied practice so that the consultant and the clinician can provide people with more accurate feedback as to their abilities. A better understanding, such as this research provides, can help ensure that test scores convey distinctive meanings that are better understood than before—a key contributor to meeting the standards of optimal assessment (Joint_Committee, 2014; Sinharay et al., 2011). It is our hope that the findings obtained here, along with further research in the area, will help promote a better understanding of the nature of the ability model of emotional intelligence and its application.

Funding

MHS of Toronto provided funding for the data collections involved in Studies 2, 3, and 4, and also provided funding to make this article open source.

Author Acknowledgments

The authors are grateful for the contributions of the staff of MHS, Toronto, past and present, to the MSCEIT, YRV, and MSCEIT-2 tests, including especially Iris Lin, Joanna Solomon, Sara Rzepa, Brady Hansma, and many others who assisted with test design and writing, data collection, and many aspects of data analysis. We acknowledge in addition the contributions of Peter Salovey who collaborated on the theoretical formulation and overall design of the tests we analyzed here.

CRediT authorship contribution statement

John D. Mayer: Conceptualization, Data curation, Formal analysis, Project administration, Software, Validation, Methodology, Writing - original draft, Writing - review & editing. **David R. Caruso:** Data curation, Formal analysis, Project administration, Software, Methodology, Writing - review & editing. **Gill Sitarenios:** Data curation, Formal analysis, Software, Validation, Methodology, Writing - review & editing. **Manolo Romero Escobar:** Data curation, Formal analysis, Software, Validation.

Declaration of competing interest

The first two authors receive royalty income from the MSCEIT, MSCEIT-YRV, and MSCEIT-2 tests from the publisher, MHS, Toronto. The third author was employed by MHS and oversaw the psychometric studies of the first two scales and most of the MSCEIT-2. The fourth author was employed by MHS, and carried out many of the analyses of the MSCEIT-2.

Data availability

In keeping with the authors' obligations, the data cannot be shared because the raw data sets are the property of the test publisher and the authors' access to them is contractually limited. The rationale for maintaining data security around the sample data is that the data involve "critical concerns regarding release of test materials primarily relate(d) to test security, the potential invalidation of tests, copyright laws, and contractual obligations." (See https://www.apa.org/science/programs/testing/data-disclosure-faqs and *Ethics Code* Standard 9.11, at https://www.apa.org/ethics/code.

Appendix A. Technical supplement

Supplementary information regarding this article can be found online at https://doi.org/10.1016/j.paid.2023.112468 and also at https://osf.io/28z5c.

References

Aiken, L. R. (1979). Relationships between the item difficulty and discrimination indexes. *Educational and Psychological Measurement, 39*(4), 821–824. https://doi.org/10.1177/001316447903900415

Bryan, V. M., & Mayer, J. D. (2021). Are people-centered intelligences psychometrically distinct from thing-centered intelligences? A meta-analysis. *Journal of Intelligence, 9* (4), 48. https://doi.org/10.3390/jintelligence9040048

Buck, R. (1984). *The communication of emotion*. Guilford Press.

Cureton, E. E. (1939). The principal compulsions of factor-analysts. *Harvard Education Review, 19*, 287–295.

Elfenbein, H. A., & MacCann, C. (2017). A closer look at ability emotional intelligence (EI): What are its component parts, and how do they relate to each other? *Social and Personality Psychology Compass. APA PsycInfo., 11*(7). https://doi.org/10.1111/spc3.12324

Ertel, S. (2011). Exploratory factor analysis revealing complex structure. *Personality and Individual Differences, 50*(2), 196–200. https://doi.org/10.1016/j.paid.2010.09.026

Fan, H., Jackson, T., Yang, X., Tang, W., & Zhang, J. (2010). The factor structure of the Mayer–Salovey–Caruso emotional intelligence test V 20 (MSCEIT): A meta-analytic structural equation modeling approach. *Personality and Individual Differences, 48*(7), 781–785. https://doi.org/10.1016/j.paid.2010.02.004

Fiori, M., & Antonakis, J. (2011). The ability model of emotional intelligence: Searching for valid measures. *Personality and Individual Differences, 50*(3), 329–334. https://doi.org/10.1016/j.paid.2010.10.010

Flora, D. B., & Flake, J. K. (2017). The purpose and practice of exploratory and confirmatory factor analysis in psychological research: Decisions for scale development and validation. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement, 49*(2), 78–88. https://doi.org/10.1037/cbs0000069

Gorsuch, R. L. (2015). *Factor analysis*. Routledge/Taylor & Francis Group.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Joint_Committee. (2014). *Standards for educational and psychological testing*. American Psychological Association.

Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology, 95*(1), 54–78. https://doi.org/10.1037/a0017286

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*, 159–174. https://doi.org/10.1016/j.intell.2014.09.005

Lewis, G. J., Lefevre, C. E., & Young, A. W. (2016). Functional architecture of visual emotion recognition ability: A latent variable approach. *Journal of Experimental Psychology. General, 145*(5), 589–602.

Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods, 18*(3), 285–300.

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), 358–374.

Markon, K. E. (2019). Bifactor and hierarchical models: Specification, inference, and interpretation. *Annual Review of Clinical Psychology, 15*, 51–69. https://doi.org/10.1146/annurev-clinpsy-050718-095522

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In D. J. Sluyter (Ed.), *Emotional development and emotional intelligence: Educational implications*. . (1–Book, Section, pp. 3–34). Basic Books.

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. (2002). *Mayer-Salovey-Caruso Emotional Intelligence Test: (MSCEIT): User's Manual*. Toronto: MHS.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion, 3*(1), 97–105. https://doi.org/10.1037/1528-3542.3.1.97

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). Emotional Intelligence: Theory, findings, and implications. *Psychological Inquiry, 15*(3), 197–215.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2014). *Mayer-Salovey-Caruso Emotional Intelligence Test: Youth Research Version*. MSCEIT: YRV. Toronto: MHS.

Mayer, J. D., Caruso, D. R., & Salovey, P. (2016). The ability model of emotional intelligence: Principles and updates. *Emotion Review, 8*, 1–11.

Mayer, J. D., Caruso, D. R., Sitarenios, G., & Romero Escobar, M. (2023). *Technical Supplement for How Many Emotional Intelligence Abilities Are There? An Examination of Four Measures of Emotional Intelligence*. https://doi.org/10.1016/j.paid.2023.112468. https://osf.io/28z5c

McGrew, K. S., Schneider, W. J., Decker, S. L., & Bulut, O. (2023). A psychometric network analysis of CHC intelligence measures: implications for research, theory, and interpretation of broad CHC scores "Beyond g.". *Journal of Intelligence, 11*(1), 19. https://doi.org/10.3390/jintelligence11010019

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49*(1), 115–132. https://doi.org/10.1007/BF02294210

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide (Eighth)*. Muthén & Muthén.

Nunnally, J. C. (1967). *Psychometric theory (2003-00036-000)*. APA PsycInfo: McGraw-Hill.

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* (3), 285–305. https://doi.org/10.1016/j.intell.2004.11.003

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

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), 196–231. APA PsycInfo https://doi.org/10.1037/1528-3542.1.3.196.

Schlegel, K., & Mortillaro, M. (2019). The Geneva Emotional Competence Test (GECo): An ability measure of workplace emotional intelligence. *Journal of Applied Psychology, 104*(4), 559–580. https://doi.org/10.1037/apl0000365

Schlegel, K., Grandjean, D., & Scherer, K. R. (2012). Emotion recognition: Unidimensional ability or a set of modality- and emotion-specific skills? *Personality and Individual Differences, 53*(1), 16–21. https://doi.org/10.1016/j.paid.2012.01.026

Schlegel, K., Boone, R. T., & Hall, J. A. (2017). Individual differences in interpersonal accuracy: A multi-level meta-analysis to assess whether judging other people is one skill or many. *Journal of Nonverbal Behavior, 41*(2), 103–137. https://doi.org/10.1007/s10919-017-0249-0

Schlegel, K., Palese, T., Mast, M. S., Rammsayer, T. H., Hall, J. A., & Murphy, N. A. (2019). A meta-analysis of the relationship between emotion recognition ability and intelligence. *Cognition and Emotion.*. https://doi.org/10.1080/02699931.2019.1632801

Schweizer, K., & Troche, S. (2018). Is the factor observed in investigations on the item-position effect actually the difficulty factor? *Educational and Psychological Measurement, 78*(1), 46–69. https://doi.org/10.1177/0013164416670711

Sinharay, S., Puhan, G., & Haberman, S. J. (2011). An NCME instructional module on subscores. *Educational Measurement: Issues and Practice, 30*(3), 29–40. https://doi.org/10.1111/j.1745-3992.2011.00208.x

Thurstone, L. L. (1935). *Isolation of primary factors. In the vectors of mind: Multiple-factor analysis for the isolation of primary traits* (pp. 171–198). University of Chicago Press. https://doi.org/10.1037/10018-007

Thurstone, L. L. (1940). Current issues in factor analysis. *Psychological Bulletin, 37*(4), 189–236. APA PsycInfo https://doi.org/10.1037/h0059402.

de Winter, J. C. F., Dodou, D., & Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research, 44*(2), 147–181. https://doi.org/10.1080/00273170902794206

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods, 51*(1), 409–428. APA PsycInfo https://doi.org/10.3758/s13428-018-1055-2.