



Differences between multimedia and text-based assessments of emotion management: An exploration with the multimedia emotion management assessment (MEMA)

Carolyn MacCann, Filip Lievens, Nele Libbrecht & Richard D. Roberts

To cite this article: Carolyn MacCann, Filip Lievens, Nele Libbrecht & Richard D. Roberts (2016) Differences between multimedia and text-based assessments of emotion management: An exploration with the multimedia emotion management assessment (MEMA), *Cognition and Emotion*, 30:7, 1317-1331, DOI: [10.1080/02699931.2015.1061482](https://doi.org/10.1080/02699931.2015.1061482)

To link to this article: <http://dx.doi.org/10.1080/02699931.2015.1061482>



Published online: 12 Aug 2015.



Submit your article to this journal [↗](#)



Article views: 151



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Differences between multimedia and text-based assessments of emotion management: An exploration with the multimedia emotion management assessment (MEMA)

Carolyn MacCann^a, Filip Lievens^b, Nele Libbrecht^b and Richard D. Roberts^c

^aSchool of Psychology, The University of Sydney, Sydney, Australia; ^bDepartment of Personnel Management, Work & Organizational Psychology, Ghent University, Ghent, Belgium; ^cProfessional Examination Service, New York, NY, USA

ABSTRACT

People process emotional information using visual, vocal, and verbal cues. However, emotion management is typically assessed with text based rather than multimedia stimuli. This study ($N=427$) presents the new multimedia emotion management assessment (MEMA) and compares it to the text-based assessment of emotion management used in the MSCEIT. The text-based and multimedia assessment showed similar levels of cognitive saturation and similar prediction of relevant criteria. Results demonstrate that the MEMA scores have equivalent evidence of validity to the text-based MSCEIT test scores, demonstrating that multimedia assessment of emotion management is viable. Furthermore, our results inform the debate as to whether cognitive saturation in emotional intelligence (EI) measures represents “noise” or “substance”. We find that cognitive ability associations with EI represent substantive variance rather than construct-irrelevant shared variance due to reading comprehension ability required for text-based items.

ARTICLE HISTORY

Received 13 November 2014
Revised 14 May 2015
Accepted 8 June 2015

KEYWORDS

Emotional intelligence;
emotion management;
situational judgment test;
multimedia assessment;
academic achievement

The Mayer–Salovey model of emotional intelligence (EI) has emotion management as its apex, with meta-analytic evidence suggesting that emotion management is the key ingredient of EI for predicting important real-world outcomes (Joseph & Newman, 2010; Mayer, Roberts, & Barsade, 2008). Existing assessments of EI measure emotion management with text-based situational judgment test (SJT) items (e.g., Freudenthaler & Neubauer, 2005; MacCann & Roberts, 2008; Mayer & Salovey, 1997; Mayer, Salovey, & Caruso, 2002). However, emotion research has established that emotions are expressed and perceived through visual and auditory cues, such as tone of voice, as well as dynamic face and body movements (e.g., Bänziger & Scherer, 2005; Sauter, Eisner, Calder, & Scott, 2010; Scherer & Ellgring, 2007; Ververidis & Kotropoulos, 2006). Moreover, criterion-related validity is higher for multimedia than for text-based SJTs in the related domain of interpersonal constructs

(meta-analytic $r = .47$ versus $.27$; Christian, Edwards, & Bradley, 2010, see also Lievens & Sackett, 2006). Taken together, these results suggest that emotion management might be more accurately measured with multimedia than text-based SJTs.

There are two objectives of the current study. The first objective is to report on the development of a multimedia SJT assessing emotion management (the multimedia emotion management assessment; MEMA). The second objective is to test the potential differences between text-based and multimedia assessments of emotion management, focusing on: (1) cognitive saturation and (2) prediction of well-being and other valued outcomes such as academic achievement. The paragraphs below describe the rationale for developing the MEMA, the process used, and our rationale for positing differences between text-based and multimedia tests of emotion management.

Assessing EI with an ability framework

The current study uses an ability framework for defining and measuring EI. Specifically, EI is defined according to Mayer, Caruso, and Salovey (1999) of four ability branches: (1) identifying emotions; (2) using emotions to facilitate thought; (3) understanding emotions; and (4) managing emotion. However, recent research suggests that the “using emotions” concept is both empirically and conceptually redundant with other branches (Fan, Jackson, Yang, Tang, & Shang, 2010; Joseph & Newman, 2010; MacCann, Joseph, Newman, & Roberts, 2014), so we do not examine this branch in the current study. Our comparison of text-based versus multimedia assessments contrasts the (text-based) MSCEIT Management assessment against the newly developed (multimedia) MEMA.¹ These measures are placed in a broad nomological network of other EI measures, intelligence, and outcome variables such as life satisfaction and academic achievement (note that the validity argument for academic achievement involves both direct prediction and indirect prediction via differential use of coping strategies).

The Mayer–Salovey–Caruso Emotional Intelligence Test

The vast majority of research on EI abilities has used the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT; Mayer et al., 2002). The MSCEIT is the only ability-based instrument to assess all four branches of EI, and includes two subtests for each branch. All of the tests measuring the use, understanding, and management of emotion are entirely text based. The two perception tests use static visual images, where test-takers rate the extent to which emotions are present in human faces or other pictures. The MSCEIT test item stimuli do not encompass the processing of tone-of-voice, dynamic facial expressions (e.g., micro-expressions), posture, or other multi-channel emotional information.

Alternatives to the MSCEIT: assessments of emotion identification

Besides the MSCEIT, there are several assessments that involve identifying emotion in a tone of voice,

facial expression, body language depiction, micro-expression, or other relevant stimuli. These include but are not limited to: the Diagnostic Analysis of Non-verbal Accuracy (DANVA; Nowicki & Carton, 1993); the Japanese and Caucasian Brief Affect Recognition Test (JACBART; Matsumoto et al., 2000); and the emotion recognition tests from the Geneva affective sciences group (Multimodal Emotion Recognition Test and Geneva Emotion Recognition Test; Bänziger, Grandjean, & Scherer, 2009; Schlegel, Grandjean, & Scherer, 2014). Research outside of the MSCEIT measurement tradition has often used the nomenclature “emotion recognition” to describe such capacities, whereas the MSCEIT refers to “emotion perception”.

Evidence to date suggests that emotion recognition tests do not assess the same underlying construct as the MSCEIT emotion perception tests (MacCann, Pearce, & Roberts, 2011; Roberts et al., 2006). There are two possible reasons for this. First, MSCEIT tests require a *quantitative* judgement (e.g., the quantity of emotion present in a facial expression), whereas emotion recognition tests usually require a *qualitative* judgement (e.g., which of many qualitatively different emotions are portrayed). Both processes are logically involved in real-world processing of emotional information—identifying both the type and extent of emotion portrayed is required to respond appropriately to an emotional situation. Second, the MSCEIT emotion perception tasks require further development, and are psychometrically weak assessments of emotion perception (Maul, 2012). Due to the possible distinction between emotion perception and emotion recognition, we include both a recognition measure (the DANVA Faces) and a perception measure (MSCEIT Perception) when comparing validity evidence for the MSCEIT Management and MEMA.²

Alternatives to the MSCEIT: higher branches of EI

For the other three branches of EI, there are fewer alternatives to the MSCEIT and *no* alternative tests that use multimedia stimuli (to our knowledge). Thus, one objective of the current study was to fill this key gap in extant EI measurement paradigms by developing a multimedia SJT of emotion management

¹Note that we chose to compare the MEMA to the most widely used emotion management test rather than to a text-based version of the MEMA. This was due to concerns with transcribing (i.e., transcriptions of verbal text versus non-verbal expressions, see Lievens & Sackett, 2006, p. 1183, see also the Discussion section below), and to ensure a comparison to an assessment intentionally designed as a text-based test.

²Note that the Geneva Emotion Recognition Task is a better representation of emotion recognition in dynamic stimuli, as compared to the DANVA. This assessment was not yet available when we collected our data.

and testing the differences between multimedia and text-based assessments.

One of the primary methods for assessing emotion management is the SJT paradigm, where test-takers must evaluate multiple possible responses to a hypothetical scenario (Libbrecht, Lievens, Carette, & Côté, 2014; MacCann & Roberts, 2008). Item types include multiple choice; select the best and worst responses; rate each option; rank-order the options; or distribute points among the options. McDaniel and Nguyen (2001) suggest a three-step methodology for developing SJT items: (1) item construction; (2) response option generation; and (3) scoring key development. Meta-analyses show that SJT scores not only correlate with personality and intelligence, but also incrementally predict job performance above personality and intelligence (Christian et al., 2010; McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001; McDaniel, Hartman, Whetzel, & Grubb, 2007).

Several research teams have developed emotion management assessments using this three-step approach. For example, MacCann and Roberts (2008) developed the Situational Test of Emotion Management (STEM). Development of the STEM used critical incidents interviews for the item and response generation and a group of experts for the scoring key development (for a detailed description, see MacCann & Roberts, 2008). Available validity evidence for these new SJT-based EI tests is promising (e.g., Burrus et al., 2012; Libbrecht & Lievens, 2012; Libbrecht et al., 2014), though ongoing research continues to be directed at improving the design of these measures and underlying scoring models (see e.g., Allen, Weissman, Hellwig, MacCann, & Roberts, 2014; Allen, Rahman, Weissman, MacCann, & Roberts, 2015).

However, no one has yet extended this SJT methodology to include multimedia stimuli in emotion management SJTs. In such multimedia SJTs, scenarios and response options are viewed in video form (Lievens, Buyse, & Sackett, 2005; Weekley & Jones, 1997). The lack of multimedia formats for assessing emotion management is surprising in light of the potential conceptual and empirical reasons as to why a multimedia format might improve EI measurement over the traditional text-based format. Below we develop such hypotheses regarding key differences between multimedia and text-based assessments of emotion management. To date, we do not know which of these reasons might explain the potential superiority of multimedia formats over text-based ones.


Development and background to the MEMA

The MEMA was developed as a multimedia adaptation of the STEM. Two professional item writers were provided with STEM items as well as transcripts of the interviews from which these items were developed. On the basis of the text-based items, scripts for short video-based scenarios and responses were developed. Both the scenarios and responses to the scenarios were presented in multimedia format, and both the scenario and responses could be played multiple times (see Figure 1 for an illustration). For all items, the scenario showed an interaction among the protagonist and at least one other person (including facial expressions, body language, and prosody vocal information) and the response options showed a single protagonist in head and shoulders profile (though hand gestures were included where relevant). Figure 2 shows an example of the original STEM item transformed into a multimedia script. Sixteen scripts were created in this way (eight from each item writer), and were reviewed (and revised accordingly) by an expert panel consisting of the STEM authors, other assessment development staff, and the director and production company staff. The MEMA also included response justification mechanisms (see Figure 2). After evaluating the scenario responses, the next page asked test-takers why they had selected that particular response, where test-takers had to choose from five options that described the general class of behaviour each response represented, including the intended goals or purpose of the response. These justification mechanisms were included to discourage random responding and to tap explicit as well as implicit knowledge about emotion management strategies.

Of the 16 scenarios (each with four possible responses), 6 were presented in multiple choice format, 6 in rank-order format (i.e., test-takers had to rank the options from best to worst), 2 asked test-takers to categorise each option as appropriate or inappropriate, and 2 asked test-takers to rate the responses on a 7-point scale “extremely bad” (1) to “extremely good” (7). This resulted in $12 + 8 + 8 = 28$ items. All items were scored dichotomously in line with expert and consensus weighting from the original STEM items. That is: (a) for *multiple-choice items*, the best option received 1 and the others zero; (b) for *rank-order items*, four items were scored correct if the worst item was ranked last, and two items were scored correct if the best item was ranked first; (c) for

Scenario

Use the PLAY button to view their dialogue on Brad's academic progress. Then click "Next" in order to see four potential responses from Brad.




Response options


Click on the **BEST** choice.

You can replay any of the videos by clicking on the Play buttons.


A



B



C



D




Figure 1. Example scenario and response options from the MEMA.

appropriate/inappropriate items, correctly specifying “appropriate” or “inappropriate” scored 1 and incorrect specification scored 0; (d) for rating-scale items, selecting the best rating scale point (or one of the two points either side) scored one. A full point was awarded for a correct answer with correct justification, and half a point was awarded for an incorrect answer but consistent justification for that response. For information about obtaining the MEMA stimuli, contact the last author.

H1: the MEMA will show adequate evidence of construct validity

We consider four key criteria for evaluating the construct validity of the MEMA. First, given that the MEMA is a test of EI, MEMA scores should show large positive correlations with existing tests of EI

(H1a). Second, given that the MEMA is intended to assess the management branch of EI, MEMA scores should show stronger correlations with emotion management than with other branches of EI (H1b). Third, in line with validity arguments that EI constitutes a group factor of intelligence (cf. MacCann et al., 2014), we expect that MEMA scores will be significantly correlated with scores on traditional tests of intelligence (H1c), but will be clearly distinct from broad personality domains (H1d), showing small to zero associations. We also consider the prediction of emotion-related criteria, and compare the MEMA with the MSCEIT Management in terms of their prediction of such outcomes. Our expectations are that the MEMA may show superior evidence of criterion-related validity for the reasons outlined below.

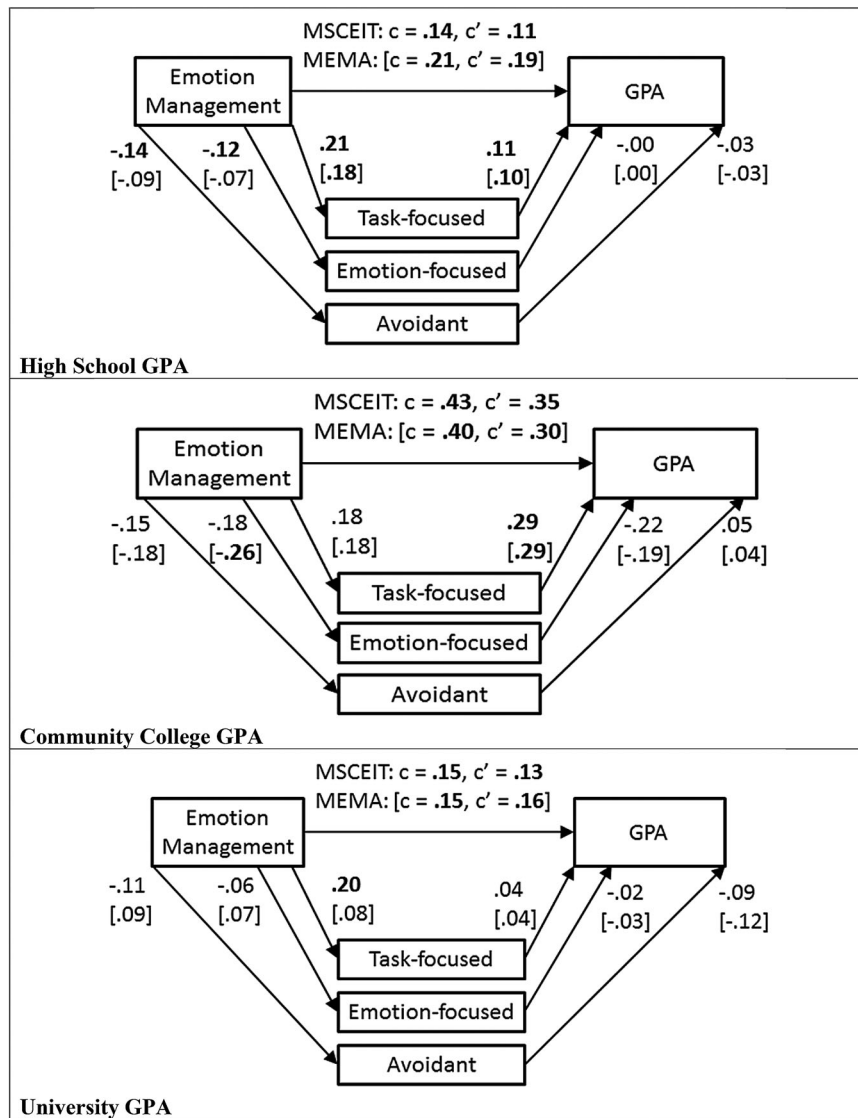


Figure 2. Mediation models testing whether coping mediates the emotion management/GPA relationship. Models were run separately for the MEMA and MSCEIT emotion management, with parameters for the MEMA shown in brackets. Parameters significant at $p < .05$ are shown in bold text.

Potential differences between multimedia and text-based assessments of emotion management

There are both conceptual and empirical reasons to expect that multimedia formats will be better than text-based formats for assessing emotion management. In the paragraphs below, we elaborate on these conceptual and empirical reasons underlying hypotheses for the differences between text-based and multimedia measurements of emotion management. We see these potential differences as

complimentary instead of as competing explanations for the superiority of multimedia measurement of emotion management.

H2: incidental cognitive saturation will be higher for text-based than multimedia format

Cognitive saturation is defined as the correlation between an instrument's test scores and performance on a measure of cognitive ability (Lubinski, 2004). In ability models of EI, emotion management is conceptualised as a cognitive ability, and should therefore

show positive relationships with other tests of cognitive ability (e.g., MacCann et al., 2014). However, the relationship with other intelligence tests should be due to the cognitive processes involved in managing emotions, rather than the ability to read the test instructions and items. Although both text-based and multimedia formats of emotion management assess the cognitive processes underlying emotion management, the text-based modality may add unintended cognitive load to the assessment because it requires participants to read and understand the item stem. So, the text-based format requires a threshold level of reading comprehension that may constitute a source of construct-irrelevant variance, leading to inflated correlations with cognitive ability. Conversely, the multimedia format for measuring emotion management does not require reading comprehension for interpreting the item stem and the response options and should therefore impose fewer unintentional cognitive requirements compared to a text-based format.

In the current study, both fluid intelligence (Gf) and crystallised intelligence (Gc) are included as aspects of cognitive ability. Gf is the ability to reason logically with novel stimuli, whereas Gc represents acculturated knowledge. Both of these components have been included in major structural models of intelligence (McGrew, 2009), and both have shown strong relationships with text-based EI assessments, including emotion management tests (e.g., MacCann, 2010; MacCann et al., 2014; Roberts, Schulze, & MacCann, 2008). At present, it is unclear whether these known relationships between EI and cognitive ability are due to the reading comprehension load of text-based assessment (i.e., represent artefacts or “noise” in the measurement of EI) or due to a genuine relationship between constructs (i.e., represent construct-relevant variance or “substance”).

H3: stronger associations with relevant criteria will be found for multimedia than for text-based formats

A multimedia presentation may enhance the criterion-related validity of emotion management scores through increased stimulus fidelity. Fidelity is the degree to which the assessment task and context mirror those actually present in the real environment and is a key determinant of validity (Callinan & Robertson, 2000). As the multimedia SJT items display a more comprehensive and richer representation of the situation, they bear a closer resemblance to real

emotional situations, which may lead to better construct representation and better prediction of criteria. Empirically, such evidence for validity differences between multimedia and text-based assessments was found in the selection literature. Results comparing multimedia versus text-based SJTs of interpersonal skills support this assertion—a multimedia format showed stronger prediction of criteria than a text-based format (Christian et al., 2010; see also Lievens & Sackett, 2006).

In this study, we consider a broad variety of criteria for comparing the validity of the multimedia versus text-based formats for assessing emotion management. We include criterion measures where relationships with EI have previously been suggested: Life satisfaction, quality of life, typical ways of coping with stress, and perceived social support (Ciarrochi, Chan, & Caputi, 2000; Lopes et al., 2004; MacCann & Roberts, 2008; MacCann, Fogarty, Zeidner, & Roberts, 2011). We also measured personality to assess whether multimedia and text-based assessments show equivalent levels of discriminant validity evidence.

We also consider the effect of emotion management on student’s school achievement, in line with emergent evidence suggesting that academic success is due to social and emotional factors as well as mastering academic material (e.g., Libbrecht et al., 2014). A recent meta-analysis showed that programmes targeting social and emotional competencies increased academic achievement by nearly a third of a standard deviation, providing evidence that emotional competencies influence academic achievement (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). One mechanism explaining why EI predicts achievement is that higher EI students may use more effective strategies for coping with school-based stressors, such that their achievement is less impeded by stress (MacCann et al., 2011). Academic success involves effectively coping with stressful tasks, including high-stakes assessments, learning from negative or critical feedback, persevering after failure, and the social demands of collaborative work (e.g., MacCann, Lipnevich, Burrus, & Roberts, 2012). In line with previous research where coping mediates the EI/achievement relationship, we test both the direct relationship between emotion management and academic achievement and the indirect relationship through coping (see Figure 2 for an illustration of the indirect relationship).

Method

Participants

Participants were 427 college students attending either community college (33%) or university (67%) in the USA (58% female). Most participants (82%) were aged under 23 years ($M = 21.6$ years, $SD = 5.9$, range between 17 and 59 years). Reported ethnicities were 69% White, 13% Black or African-American, 10% Hispanic, with the remaining 8% Asian, Native American, multiple ethnicities, or "other".

Procedure

Testing sessions took place at college or university computer laboratories, proctored by college/university staff. Two counter-balanced test orders were randomly assigned to participants. To avoid fatigue, participants could respond over two sessions, with each session containing two breaks (average time interval between these sessions = one week). The assessment battery was composed of five EI tests, a comprehensive set of intelligence tests, a personality measure, and several criterion measures. Students were compensated for their participation. Subsets of these data have been used in previous research (Legree, Pstoka, Robbins, Roberts, & Putka, 2014; MacCann et al., 2011, 2014). The current data-set only included participants who had completed the MEMA. No other exclusion criteria were used.

Predictor measures

MEMA; Roberts et al. (2011). Participants completed the 28-item MEMA (as described in the introduction; see Table 1 for an example item).

Emotion recognition. The Diagnostic Analysis of Nonverbal Accuracy-2, Adult Facial Expressions (DANVA-2-AF; Nowicki & Carton, 1993) consists of 24 photographs of adults who express happiness, sadness, anger, or fear, in either high or low intensity. Pictures were displayed for two seconds and respondents indicated which of the four emotions was present in each photograph.

Performance-based emotional intelligence test. The MSCEIT (Mayer et al., 2002) contains 141 items spread over 8 subtests, 2 for each of the 4 EI branches. To measure *emotional perception* respondents were asked to identify emotions in faces and in pictures. To assess *emotional facilitation*, respondents were

asked to identify which emotions promote specific thoughts/activities, and which combination of sensations is similar to three different emotions. *Emotional understanding* was measured by asking respondents to define which emotions arise from different emotion blends, and to identify transitions between emotions. Finally, *emotional management* was assessed by presenting respondents with vignettes describing a social situation and asking them to rate the effectiveness of different strategies to manage emotions of self (emotional management) and others (emotional relationships). The MSCEIT was scored according to consensus scoring.

Cognitive abilities. Intelligence was measured with six cognitive tests from the Kit of Factor Referenced Cognitive Tests developed by ETS (Ekstrom, French, Harman, & Derman, 1976). Three tests were used to measure fluid ability (letters, figure classification, and calendar), and three tests were used to measure crystallised ability (vocabulary, analogies, and sentences). In the *letters* test, respondents have to identify which set of letters does not fit the rule that makes the other sets of letters alike (15 items). In the *figure classification* test, respondents must indicate to which group a figure belongs (20 items). In the *calendar* test, respondents are presented with a calendar and have to select the correct answer to date-related questions about this calendar (20 items). In the *vocabulary* test, respondents have to select the word most similar to a target word from 4 alternatives (36 items). In the *analogies* test, respondents have to find the relationship between a pair of given words and identify the pair of words from the answer choices with the same relationship (30 items). In the *sentence completion* test, respondents are presented with sentences containing one or two blanks and have to select the best answer choice to make that sentence make sense (30 items).

Personality. Apart from cognitive abilities, we also measured personality to assess evidence of whether the MEMA had discriminant validity in this area. Personality traits were assessed with 40 items for each of the Big Five personality traits with 5-point Likert scales, ranging from 1 (not at all like me) to 5 (very much like me). Items were drawn from the International Personality Item Pool (Goldberg, 1999).

Criterion measures

Life satisfaction. Life satisfaction was assessed using the seven-item Students' Life Satisfaction Scale (Huebner, 1991). Respondents used a 6-point Likert

Table 1. Illustration of how the original text-based STEM item was transformed into a script for filming of the MEMA.

STEM item	Script for MEMA item
<p><i>Scenario</i> Mario is showing Min, a new employee, how the system works. Mario's boss walks by and announces Mario is wrong about several points, as changes have been made. Mario gets on well with his boss, although they do not normally have much to do with each other</p>	<p><i>Scenario (video: 45sec)</i> <i>Maria (a Hispanic female, age unimportant) is standing with James (an Asian-American male, age unimportant) in front of a filing cabinet</i> Maria [to James]: So it's really pretty straightforward. As soon as the invoices are sent down to Laura, just create a new red folder and file them by invoice number. Make sure you also create a green file and put it under the customer's name at the same time, because if you forget about it, that can cause lots of problems down the road JAMES [a little overwhelmed]: So, wait—the RED file is for the new invoices? Aren't some of the invoice files yellow? <i>Angela (a white woman, mid-forties to mid-fifties) overhears the conversation as she walks by. She stops, listens as James finishes his question, and interrupts before Maria can respond</i> ANGELA: Maria, you're explaining it all wrong—weren't you at the meeting last week? We've changed the color-coding system. New invoices are yellow now, not red. You're just going to confuse him!</p>
<p>What action would be the most effective for Mario?</p>	<p><i>Response options (video: 15–30second per response)</i> A. MARIA (humorously): Well, you certainly seem to know your stuff, Angela. You're a virtual color wheel! B. MARIA (angrily): Angela, I'm trying to explain the system to James, and you're just confusing him. Why don't you go work with your own intern and leave us to this? C. MARIA (calmly): Actually, I did miss that meeting last week. James, let's go find Lana and make sure we're doing this right* D. MARIA (apologetically): I'm really sorry, James; I think I messed up. Just let Angela tell you how to do the files; she seems to know more about them than I do</p>
<p><i>Response options</i> A. Make a joke to Min, explaining he didn't know about the changes B. Not worry about it, just ignore the interruption C. Learn the new changes* D. Tell the boss that such criticism was inappropriate</p>	<p><i>Response justifications (text). Why did you choose this as the best response? Select all that apply</i> Maria directly expresses concern to James for her mistake. (Justifies C or D) Maria shows Angela that her interruption is rude and distracting. (Justifies B) Maria admits that she might have made an error. (Justifies C or D) Maria attempts to find correct information for James. (Justifies C)</p>

Note: The best item is marked with an asterisk for both STEM and MEMA.

scale, ranging from 1 (*strongly disagree*) to 6 (*strongly agree*).

Quality of life. Quality of life was measured with the 24-item World Health Organization Quality of Life—Abbreviated Version (Skevington, Lotfy, & O'Connell, 2004). The questionnaire contained assesses four quality of life domains (physical, psychological, social, and environmental). Respondents used a 5-point scale, ranging from 1 (*not at all*) to 5 (*extremely*) to indicate their response to each item.

Coping. Students' coping behaviour was measured using the 24-item Coping with School Situations (MacCann et al., 2011). This scale consists of three scales assessing task-oriented coping, emotion-oriented coping, and avoidant coping. Respondents rated how each item on a 5-point scale ranging from 1 (*almost never*) to 5 (*almost always*).

Perceived social support. The Multidimensional Scale of Perceived Social Support (Zimet, Dahlem, Zimet, & Farley, 1988) is 12-item scale designed to measure

perceived social support from friends, family, and significant others. The respondents rated their agreement on a 7-point scale, ranging from 1 (*very strongly disagree*) to 7 (*very strongly agree*).

GPA. Participants reported their high school GPA and their GPA at community college (2 years) or university (4 years). All GPA was reported on a scale on 0–4. Prior research has found high convergence between self-reported GPA and actual GPA (Kuncel, Crede, & Thomas, 2005). GPA was analysed separately for community college versus four-year college students, as these two educational systems serve different functions and different populations, and often have markedly different standards for grading.

Results

Table 2 presents the reliability, descriptive statistics, and mean gender differences for all study variables. In addition, Table 2 presents the correlation of the

Table 2. Descriptive statistics and correlations with MSCEIT Management and MEMA scores for all study variables ($N = 394$; missing data treated list-wise).

	Descriptive statistics							<i>d</i>	Correlations	
	All			Male		Female			MSCEIT manage	MEMA
	<i>a</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Gc										
Vocabulary	0.82	0.62	0.16	0.61	0.17	0.61	0.16	0.00	0.40**	0.46**
Analogies	0.80	0.47	0.18	0.48	0.21	0.46	0.16	0.09	0.44**	0.45**
Sentences	0.90	0.86	0.17	0.85	0.19	0.87	0.16	-0.10	0.53**	0.55**
Gf										
Letters	0.80	0.68	0.23	0.65	0.24	0.69	0.23	-0.15	0.46**	0.39**
Figures	0.91	0.68	0.13	0.67	0.13	0.69	0.12	-0.18	0.47**	0.41**
Calendar	0.84	0.62	0.22	0.59	0.24	0.63	0.22	-0.20	0.53**	0.45**
Personality										
Openness	0.92	3.57	0.51	3.51	0.49	3.62	0.52	-0.22*	0.29**	0.33**
Conscientiousness	0.90	3.40	0.46	3.33	0.43	3.46	0.46	-0.30**	0.22**	0.13**
Extraversion	0.91	3.35	0.48	3.35	0.45	3.36	0.50	-0.02	0.04	0.08
Agreeableness	0.84	3.57	0.37	3.37	0.32	3.70	0.33	-1.03**	0.33**	0.29**
Neuroticism	0.92	2.57	0.52	2.47	0.50	2.63	0.52	-0.31**	-0.07	-0.10*
EI										
MEMA	0.61	0.76	0.10	0.73	0.12	0.77	0.09	-0.36**	0.44**	-
MSCEIT total	0.95	0.44	0.09	0.41	0.10	0.47	0.08	-0.60**	0.85*** ^a	0.53**
Perception	0.90	0.52	0.11	0.49	0.11	0.54	0.10	-0.41**	0.43*** ^a	0.31**
Understanding	0.88	0.46	0.12	0.43	0.13	0.48	0.10	-0.46**	0.72*** ^a	0.53**
Management	0.88	0.36	0.11	0.32	0.11	0.38	0.11	-0.54**	-	0.44**
DANVA-2	0.48	0.81	0.10	0.79	0.10	0.83	0.09	-0.47**	0.21*** ^a	0.34**
Outcomes										
Life satisfaction	0.84	4.50	0.93	4.47	0.90	4.52	0.95	-0.05	0.17**	0.16**
Quality of life	0.87	3.82	0.47	3.84	0.46	3.81	0.48	0.05	0.08	0.12**
Task coping	0.73	3.52	0.57	3.42	0.54	3.59	0.58	-0.31**	0.20**	0.18**
Avoidant coping	0.88	2.60	0.78	2.72	0.79	2.52	0.76	0.26**	-0.15**	-0.07
Emot coping	0.72	2.57	0.60	2.61	0.61	2.54	0.60	0.11	-0.11*	-0.09
MPSS	0.92	5.62	1.09	5.46	1.09	5.77	1.06	-0.29**	0.24**	0.25**
GPA										
High school ($n = 424$)		3.35	0.67	3.21	0.68	3.45	0.65	-0.35**	0.14**	0.21**
CommColl ($n = 101$)		3.20	0.53	3.17	0.50	3.23	0.57	-0.11	0.43**	0.40**
University ($n = 248$)		3.18	0.54	3.10	0.55	3.22	0.52	-0.23	0.15*	0.15*

Note: Cognitive ability tests were timed and alpha reliability was computed for the subset that completed all items (N ranges from 261 for calendar test to 409 for vocabulary test).

^aCorrelations for text-based versus multimedia assessments are significantly different at $p < .05$, using Steiger's (1980) test of the difference between two dependent correlations.

* $p < .05$.

** $p < .01$.

MEMA and the MSCEIT Management branch with intelligence, personality, and criterion variables. Steiger's (1980) test of the difference between two dependent correlations assessed whether the MEMA and MSCEIT showed significantly different associations with these variables.

Descriptive statistics, reliability, and gender differences

Emotional intelligence. Internal consistency reliability of the MEMA scores was .61. However, the reliability of SJTs is often better assessed by the test-retest reliability index than internal consistency reliability due to multidimensionality (Whetzel & McDaniel, 2009). To assess test-retest reliability, we administered

the MEMA to a subset of the current sample ($n = 131$) 6 months after the first administration. Test-retest reliability was $r = .54$ and Cronbach's alpha at time 2 was .69. While relatively low, this test-retest reliability is similar previously reported results for text-based assessment of emotion management ($r = .61$ for MEIS emotion management; Mayer, Salovey, & Caruso, 2008). Internal consistency of the MSCEIT total score and branch scores was good. Internal consistency reliability of the DANVA scores was lower than has been previously reported (Cronbach's $\alpha = .48$; Nowicki & Carton, 1993; Rubin, Munz, & Bommer, 2005). On average, females scored higher than males on all EI measures, with moderate effect size.

Intelligence, personality, and criterion variables. Reliability estimates for the intelligence assessments

ranged from .80 to .91. Gender differences on intelligence were all of trivial size. Cronbach's alpha ranged from .84 to .92 across the five personality domains. Sex differences in personality were consistent with prior research (Schmitt, Realo, Voracek, & Allik, 2008). Reliability estimates for the criterion measures ranged from .72 (for emotion-focused coping) to .92 (for perceived social support). Average GPA was similar for high school, community college, and universities, and was above 3.0 out of 4.0 in all three cases.

Tests of the study hypotheses

H1: The MEMA will show adequate evidence of construct validity. Hypothesis 1 was generally supported, with three out of the four sub-hypotheses consistent with our expectations. First, the MEMA showed significant positive correlations with all branches of the MSCEIT, MSCEIT total scores, and the DANVA test of emotion recognition (in line with hypotheses). However, MEMA scores were more strongly correlated with MSCEIT understanding than with MSCEIT Management (contrary to hypotheses). Note, however, that the MSCEIT understanding is the only one of the MSCEIT's branches to be assessed using multiple choice rather than a ratings-based scale of evaluation, such that results might be interpreted as potentially method based rather than substantive. Third, the MEMA showed significantly positive correlations with all six assessments of cognitive ability, demonstrating its status as an intelligence. Fourth, MEMA scores were relatively independent from personality, correlating at around .30 with Openness and Agreeableness. This is consistent with prior research on emotion management (Joseph & Newman, 2010; MacCann, 2010).

H2: Less cognitive saturation for MEMA than MSCEIT. Hypothesis 2 was not supported. As given in Table 2, relationships with the six intelligence tests were not significantly different for the MSCEIT Management ($r = .40-.53$) versus the MEMA ($r = .39-.55$). Thus, lower cognitive saturation was ruled out as a difference between multimedia and text-based assessments of emotion management. Relationships with personality were also not significantly different for the MSCEIT versus the MEMA.

H3: Higher criterion-related validity for MEMA versus MSCEIT. Hypothesis 3 was not supported. Correlations with emotion-related criterion variables were not significantly different for the MEMA versus the MSCEIT Management branch. Correlations with academic achievement were also not significantly different for

MEMA (.14-.39) and the MSCEIT Management (.15-.43). As given in Table 2, there were small positive associations of emotion management with high school GPA and with university GPA, and moderate to large positive associations with community college GPA.

MSCEIT Management showed significantly stronger prediction of MSCEIT Perception and Understanding. However, this may be interpreted as shared method variance rather than the text-based MSCEIT Management showing greater evidence of construct validity. In support of this idea, the MEMA showed a significantly stronger relationship with the DANVA assessment (the only EI criterion measure that was not part of the MSCEIT test battery).

We tested whether the emotion management/GPA relationships were significantly mediated by the three coping strategies (task focused, emotion focused, and avoidant coping). We conducted separate analyses for the three sets of grades (high school, community college, and university), and for the MSCEIT Management versus MEMA scores. Preacher and Hayes' (2008) INDIRECT SPSS macro was used to calculate both the point estimate and the 90% bias-corrected confidence intervals for the indirect effect (see Table 3). As with the zero-order correlations, there were similar results for the MEMA and the MSCEIT. The indirect effect was significant for community college and high school GPA but not for university GPA. There was a medium effect for community college GPA and a small effect for high school GPA (cf. Shrout, & Bolger, 2002). Specifically, the indirect effect of the MEMA was .019 (*ns*), .082 ($p < .05$), and .027 ($p < .05$) for university, community college, and high school GPA, respectively. The indirect effect of MSCEIT Management was $-.009$ (*ns*), .095 ($p < .05$), and .021 ($p < .05$). More detailed results of this analysis, including parameter estimates, are given in Figure 2.

Discussion

Existing emotion management tests typically use static representations of emotional expressions or emotional situations, and disregard the dynamic nature of emotion communication processes. The goals of the current study were to develop a multimedia assessment of emotion management and to compare scores obtained from this new assessment with a typical text-based assessment. From a methodological point of view, our results demonstrate that it is possible to develop a multimedia SJT for

Table 3. Standardised estimates for coping as a mediator of the emotion management/academic achievement relationship for high school, community college, and university grades: MEMA versus MSCEIT Management.

	MEMA			MSCEIT Management		
	Point estimate	Product of coefficients	95% CI	Point estimate	Product of coefficients	95% CI
University (<i>n</i> = 245)						
Problem focused	0.008	0.570	−0.013, 0.035	0.004	0.594	−0.003, 0.022
Emotion focused	0.001	0.258	−0.006, 0.020	−0.002	−0.319	−0.022, 0.005
Avoidant	0.010	0.976	−0.002, 0.041	−0.011	−1.052	−0.041, 0.000
Total	0.019	1.245	−0.005, 0.053	−0.009	−0.690	−0.037, 0.011
Community college (<i>n</i> = 99)						
Problem focused	0.051	1.573	0.000, 0.148	0.052	1.571	0.007, 0.134
Emotion focused	0.039	1.358	−0.003, 0.142	0.050	1.461	−0.001, 0.166
Avoidant	−0.008	−0.469	−0.064, 0.016	−0.008	−0.380	−0.060, 0.024
Total	0.082	2.063*	0.014, 0.202	0.095	2.193*	0.026, 0.216
High school (<i>n</i> = 415)						
Problem focused	0.023	1.942	0.006, 0.051	0.018	1.763	0.003, 0.040
Emotion focused	0.001	0.077	−0.015, 0.014	0.000	0.093	−0.008, 0.012
Avoidant	0.004	0.143	−0.009, 0.023	0.002	0.438	−0.005, 0.016
Total	0.027	2.113*	0.009, 0.055	0.021	1.972*	0.005, 0.041

Note: Bias-corrected confidence intervals were produced using 2000 bootstrap samples.

**p* < .05.

assessing emotion management, thereby extending extant EI measurement paradigms. Moreover, from a conceptual perspective, our comparison between a multimedia format and a text-based format for measuring emotion management also advances our understanding of emotion management and its underlying processes. The paragraphs below detail the knowledge we gained related to these two goals of the study.

Construct validity evidence for the MEMA

Our results showed that the MEMA meets three of the four criteria that Orchard et al. (2009) specified for evaluating the validity evidence for a new EI test: (1) the MEMA was positively related with established intelligences; (2) the MEMA was distinct from personality traits; and (3) the MEMA related to indicators of emotional functioning and academic achievement. However, the fourth criterion that EI tests should show stronger relations to other EI tests than to cognitive ability tests was not met. It is clear that the MEMA measures an ability, but it is less clear whether this ability is part of *EI*. However, the MSCEIT Management branch showed similarly high levels of relationship to intelligence in this study, suggesting that the cognitive saturation of emotion management is not limited to the MEMA but may be a general characteristic of emotion management. That is, people need to be intelligent to manage their emotions well.

Importantly, these results inform the debate as to whether cognitive saturation in EI measurement

represents “noise” or “substance” (e.g., MacCann, 2010; Orchard et al., 2009). Prior research was not able to answer this question, as it primarily relied on text-based measures of emotion management. A key conclusion of this study is that regardless of how EI is measured, the cognitive saturation remains substantial. This speaks more in favour of cognitive saturation representing substantive variance than being a measurement artefact. Such an interpretation is consistent with the idea that EI may in fact be a component of intelligence rather than a separate construct. Recent research has modelled EI as a group factor of intelligence using text-based assessments, with equivalent status to fluid intelligence or broad visualisation abilities (MacCann, 2010; MacCann et al., 2014). Current results show that EI is an intelligence even when it is assessed with multimedia rather than text-based assessments.

Comparing MEMA with a lower fidelity task

Our results demonstrate that there are no major differences between multimedia and text-based assessments—validity evidence is similar in terms of cognitive saturation, independence from personality, and prediction of emotion-related criteria. In fact, the only significant differences between the two emotion management assessments were their relationship with other EI measures. Compared to the MEMA, MSCEIT Management showed a significantly stronger relationship with the other MSCEIT subtests. However, the non-MSCEIT EI criterion

measure (the DANVA faces test) showed a significantly stronger association with the MEMA than with the MSCEIT. Taken together, results indicate equivalent evidence of validity for the MEMA versus the MSCEIT.

Given the demonstration that multimedia clips may be used to assess the emotion management branch of EI, one intriguing direction for future research would be to explicitly assess emotion perception, understanding, and management capacities using the same multimedia clips and SJT methodology. That is, for each situation, respondents could first be asked which emotions they perceived (*emotion perception*), then asked why they believed these occurred (*emotion understanding*), and finally asked to evaluate different strategies for managing these emotional situations (*emotion management*). Both the original hierarchical model of EI and the cascading model of EI recently proposed by Joseph and Newman (2010) describe a causal relationship among the branches of EI where perception and understanding of emotions precede emotion management (Mayer et al., 1999). That is, the correct response to a situation (i.e., emotion management) requires that the situation is correctly understood, which requires that the emotional information present is correctly perceived. The within-person design afforded by using the same stimuli to assess perception, understanding, and management would allow a more explicit test of the causal links in the hierarchical and cascading models of EI at the within-person level (see Rockstuhl, Ang, Ng, Lievens, & Van Dyne, 2015, for a recent example of this approach in the domain of cross-cultural intelligence). To our knowledge, these relationships have thus far been tested exclusively at the between-person level (i.e., differences in people's emotion perception capacity predict differences in people's emotion management capacity), but are often described and conceptualised at the within-person level (e.g., differences in how well an individual perceives emotions in different situations may predict differences in how well the person can manage these different situations). A within-person approach is typical across many areas of emotion research but has not been used to test the theoretical models underpinning EI.

Limitations and future directions

In this study's comparison between formats, we tried to hold the construct (emotion management)

constant, while varying the measurement approach (text based versus multimedia). Note, though, that the tests (MEMA and MSCEIT) also differed, making the comparison less strict. An alternative possibility would have been to transcribe the MEMA scenarios and item options and administer them also in a text format (apart from the multimedia format). As noted earlier, we did not engage in this transcription for two key reasons. First, we wanted to base our comparison on the most widely used assessment of emotion management (MSCEIT emotion management measure) because this would represent the most informative comparison for the field. Second, transcribing audiovisual material is not without problems. As noted by Lievens and Sackett (2006, p. 1183), a key issue is whether one describes only verbal content or also non-verbal content. If one transcribes only verbal content, the similarity of the two measures might be troublesome because the available material to participants is not the same in both versions. If one decides to transcribe both verbal and non-verbal content, the issue then becomes how accurate the non-verbal transcriptions will be and how granular they should be. In some cases, such a transcribed version might even present more information than the multimedia version. In addition, in the text-based version, participants are then no longer required to infer the non-verbal cues themselves. All of these shows that transcribing a multimedia test to a text-based one does not solve the issue of keeping the content/construct to be measured constant.

Finally, it would be beneficial to integrate notions from emotion research into EI assessment approaches. For example, there are now emotion recognition assessments and corpora encompassing a wide range of stimuli based on facial muscle movements, body language or posture, and auditory characteristics of the human voice (e.g., Bänziger et al., 2009; Bänziger, Mortillaro, & Scherer, 2012). This type of information and stimuli has not yet been applied to assess emotion understanding and management, even though these may constitute measurable sources of information about the emotional content and loading of those situations. By coding emotional situations on these different components it would be possible to create multimedia EI tests that assess the higher branches, potentially with items of varying levels of difficulty (see also Côte, 2010). These items would then be based on critical incidents and theories about emotion recognition (in speech

and faces) providing a practical and theoretical framework.

Conclusions

This study demonstrated validity evidence for assessing emotion management using multimedia stimuli. Validity evidence was equivalent to standard text-based assessment techniques, demonstrating that multimedia presentation of emotion stimuli is a viable format. The similar cognitive saturation of multimedia versus text-based assessment is encouraging for the existing corpus of research on EI, as it suggests that the status of EI as a group factor of intelligence is not due merely to the threshold level of text comprehension requires by most existing assessment techniques (which predominantly use text-based assessment).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This was supported by the US Army Research Institute under grant W91WAW-07-C-0025

References

- Allen, V. D., Rahman, N., Weissman, A., MacCann, C., & Roberts, R. D. (2015). 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.
- 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.
- Bänziger, T., Grandjean, D., & Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: The multimodal emotion recognition test (MERT). *Emotion, 9*, 691–704.
- Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the Geneva multimodal expression corpus for experimental research on emotion perception. *Emotion, 12*, 1161–1179.
- Bänziger, T., & Scherer, K. (2005). The role of intonation in emotional expressions. *Speech Communication, 46*, 252–267.
- Burrus, J., Betnacourt, A., Holtzman, S., Minsky, J., MacCann, C., & Roberts, R. D. (2012). Emotional intelligence relates to well-being: Evidence from the situational judgment test of emotional management. *Applied Psychology: Health and Well-Being, 4*, 151–166.
- Callinan, M., & Robertson, I. T. (2000). Work sample testing. *International Journal of Selection and Assessment, 8*, 248–260.
- Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology, 63*, 83–117.
- Ciarrochi, J., Chan, A. Y. C., & Caputi, P. (2000). A critical evaluation of the emotional intelligence construct. *Personality and Individual Differences, 28*, 539–561.
- Côté, S. (2010). Taking the “intelligence” in emotional intelligence seriously. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 3*, 127–130.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students’ social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Derman, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Fan, H., Jackson, T., Yang, X., Tang, W., & Shang, J. (2010). The factor structure of the Mayer-Salovey-Caruso emotional intelligence test V2.0: A meta-analytic structural equation modeling approach. *Personality and Individual Differences, 48*, 781–785.
- Freudenthaler, H. H., & Neubauer, A. C. (2005). Emotional intelligence: The convergent and discriminant validities of intra- and interpersonal emotional abilities. *Personality and Individual Differences, 39*, 569–579.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (pp. 7–28). Tilburg: Tilburg University Press.
- Huebner, E. S. (1991). Initial development of the student’s life satisfaction scale. *School Psychology International, 12*, 231–240.
- Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology, 95*, 54–78.
- Kuncel, N., Crede, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*, 63–82.
- Legree, P. J., Pstoka, J., Robbins, J., Roberts, R. D., & Putka, D. J. (2014). Profile similarity metrics as an alternate framework to score rating-based tests: MSCEIT reanalyses. *Intelligence, 47*, 159–174.
- Libbrecht, N., & Lievens, F. (2012). Validity evidence for the situational judgment test paradigm in emotional intelligence measurement. *International Journal of Psychology, 47*, 438–447.
- Libbrecht, N., Lievens, F., Carette, B., & Côté, S. C. (2014). Emotional intelligence predicts success in medical school. *Emotion, 14*, 64–73.
- Lievens, F., & Sackett, P. R. (2006). Video-based versus written situational judgment tests: A comparison in terms of predictive validity. *Journal of Applied Psychology, 91*, 1181–1188.
- Lievens, F., Buyse, T., & Sackett, P. R. (2005). The operational validity of a video-based situational judgment test for medical college admissions: Illustrating the importance of matching predictor and criterion construct domains. *Journal of Applied Psychology, 90*, 442–452.
- Lopes, P. N., Brackett, M. A., Nezlek, J. B., Schütz, A., Sellin, I., & Salovey, P. (2004). Emotional intelligence and social

- interaction. *Personality and Social Psychology Bulletin*, 30, 1018–1034.
- Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "General intelligence", objectively determined and measured". *Journal of Personality and Social Psychology*, 86, 96–111.
- MacCann, C. (2010). Further examination of emotional intelligence as a standard intelligence: A latent variable analysis of fluid intelligence, crystallized intelligence, and emotional intelligence. *Personality and Individual Differences*, 49, 490–496.
- MacCann, C., & Roberts, R. D. (2008). New paradigms for assessing emotional intelligence: Theory and data. *Emotion*, 8, 540–551.
- MacCann, C., Fogarty, G. J., Zeidner, M., & Roberts, R. D. (2011). Coping mediates the relationship between emotional intelligence (EI) and academic achievement. *Contemporary Educational Psychology*, 36, 60–70.
- 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, 358–374.
- MacCann, C., Lipnevich, A. A., Burrus, J. B., & Roberts, R. D. (2012). The best years of our lives? Coping with stress predicts school grades, life satisfaction, and feelings about high school. *Learning and Individual Differences*, 22, 235–241.
- MacCann, C., Pearce, N., & Roberts, R. D. (2011). Emotional intelligence as assessed by situational judgment and emotion recognition tests: Building the nomological net. *Psychological Topics*, 20, 393–412.
- Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Rarogue, J., Kooken, K., Ekman, P., ... Goh, A. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behaviour*, 24, 179–209.
- Maul, A. (2012). The validity of the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) as a measure of emotional intelligence. *Emotion Review*, 4, 394–402.
- Mayer, J. D., Caruso, D., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, 27, 267–298.
- Mayer, J., Roberts, R. D., & Barsade, S. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology*, 59, 507–536.
- Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey, & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3–34). New York, NY: Basic Books.
- Mayer, J. D., Salovey, P., & Caruso, D. (2002). *The Mayer, Salovey, and Caruso Emotional Intelligence Test: Technical manual*. Toronto: Multi-Health Systems.
- Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). What is emotional intelligence and what does it predict? In P. Kyllonen, R. Roberts, & L. Stankov (Eds.), *Extending intelligence* (pp. 319–348). Mahwah, NJ: Lawrence Erlbaum.
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L., III. (2007). Situational judgment tests, response instructions and validity: A meta-analysis. *Personnel Psychology*, 60, 63–91.
- McDaniel M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Predicting job performance using situational judgment tests: A clarification of the literature. *Journal of Applied Psychology*, 86, 730–740.
- McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment*, 9, 103–113.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37, 1–10.
- Nowicki, S., & Carton, J. (1993). The measurement of emotional intensity from facial expressions. *The Journal of Social Psychology*, 133, 749–750.
- Orchard, B., MacCann, C., Schulze, R., Matthews, G., Zeidner, M., & Roberts, R. D. (2009). New directions and alternative approaches to the measurement of emotional intelligence. In C. Stough, D. Saklofske & J. D. A. Parker (Eds.), *Advances in the measurement of emotional intelligence* (pp. 321–344). New York, NY: Springer Publishing.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40, 879–891.
- Roberts, R. D., Burrus, J., Betancourt, A. C., Holtzman, S., Libbrecht, N., MacCann, C., ... Schulze, R. (2011). *Multimedia assessment of emotional abilities: Development and validation*. Unpublished technical report and data. Princeton, NJ: Educational Testing Service.
- Roberts, R. D., & Schulze, R., & MacCann, C. (2008). The measurement of emotional intelligence: A decade of progress? In G. Boyle, G. Matthews, & D. Saklofske (Eds.), *The Sage handbook of personality theory and assessment* (pp. 461–482). New York, NY: Sage.
- Roberts, R. D., Schulze, R., O'Brien, K., MacCann, C., Reid, J., & Maul, A. (2006). Exploring the validity of the Mayer-Salovey-Caruso emotional intelligence test (MSCEIT) with established emotions measures. *Emotion*, 6, 663–669.
- Rockstuhl, T., Ang, S., Ng, K. Y., Lievens, F., & Van Dyne, L. (2015). Putting judging situations into situational judgment tests: Evidence from intercultural multimedia SJTs. *Journal of Applied Psychology*, 100, 464–480.
- Rubin, R. S., Munz, D. C., & Bommer, W. H. (2005). Leading from within: The effects of emotion recognition and personality on transformational leadership behavior. *Academy of Management Journal*, 48, 845–858.
- Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. (2010). Perceptual cues in nonverbal vocal expressions of emotion. *The Quarterly Journal of Experimental Psychology*, 63, 2251–2272.
- Scherer, K. R., & Ellgring, H. (2007). Multimodal expression of emotion: Affect programs or componential appraisal patterns? *Emotion*, 7, 158–171.
- Schlegel, K., Grandjean, D., & Scherer, K. R. (2014). Introducing the Geneva emotion recognition test: An example of Rasch-based test development. *Psychological Assessment*, 26, 666–672.
- Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in big five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, 94, 168–182.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7, 422–445.
- Skevington, S. M., Lotfy, M., & O'Connell, K. A. (2004). The world health organization's WHOQOL-BREF quality of life

- assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL group. *Quality of life Research*, 13, 299–310.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251.
- Ververidis, D., & Kotropoulos, C. (2006). Emotional speech recognition: Resources, features, methods, and applications. *Speech Communication*, 48, 1162–1181.
- Weekley, J. A., & Jones, C. (1997). Video-based situational testing. *Personnel Psychology*, 50, 25–49.
- Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review*, 19, 188–202.
- Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment*, 52, 30–41.