Article

Validity of the Reading the Mind in the Eyes Test in a Brazilian Sample[1]

Fabiano Koich Miguel[2]
Universidade Estadual de Londrina, Londrina-PR, Brazil

Raissa Barquete Caramanico
Universidade Estadual de Londrina, Londrina-PR, Brazil

Eduardo Yudi Huss
Universidade Estadual de Londrina, Londrina-PR, Brazil

Ana Carolina Zuanazzi
Universidade São Francisco, Itatiba-SP, Brazil

Abstract: The Reading the Mind in the Eyes Test (RMET) is internationally used to assess emotional perception, but there are few validity studies with Brazilian samples. The test was answered by 1440 participants, along with the Computerized Test of Primary Emotions Perception (PEP), and abstract (AR) and verbal reasoning (VR) tasks. RMET items were studied with Rasch model. Results indicate that its items are concentrated at a lower level of difficulty, lacking difficult items to assess higher levels of emotional perception. Both RMET and PEP showed significant correlations with AR and VR, corroborating other studies showing emotional perception is related to other types of intelligence. However the correlation between RMET and PEP was lower than expected ($r = .43$), suggesting perception of emotions in the eyes is only partially related to perception in the whole face.

Keywords: facial expressions, intelligence, psychological assessment, item response theory

Validade do Reading the Mind in the Eyes Test em uma Amostra Brasileira

Resumo: O Reading the Mind in the Eyes Test (RMET) é utilizado internacionalmente para avaliação da percepção emocional, mas são poucos os estudos de validade com amostras brasileiras. O teste foi respondido por 1440 participantes, juntamente com o Teste Informatizado de Percepção de Emoções Primárias (PEP) e provas de raciocínio abstrato (RA) e verbal (RV). Os itens do RMET foram estudados com modelo de Rasch. Os resultados indicaram que os itens estão concentrados em um nível menor de dificuldade, com falta de itens difíceis para avaliar níveis mais altos de percepção emocional. Tanto o RMET quanto o PEP mostraram correlações significativas com RA e RV, corroborando estudos que mostram que percepção emocional está relacionada a outros tipos de inteligência. Contudo a correlação entre RMET e PEP foi menor do que esperada ($r = .43$), sugerindo que percepção de emoção nos olhos está apenas parcialmente relacionada a percepção na face inteira.

Palavras-chave: expressões faciais, inteligência, avaliação psicológica, teoria de resposta ao item

Validez de la Reading the Mind in the Eyes Test en una Muestra Brasileña

Resumen: La Reading the Mind in the Eyes Test (RMET) se utiliza internacionalmente para evaluar la percepción emocional, pero son pocos los estudios de validez con muestras brasileñas. La prueba fue respondida por 1440 participantes, junto con la Prueba Computarizada de Percepción de Emociones Primarias (PEP), y tareas de razonamiento abstracto (RA) y verbal (RV). Los ítems de la RMET fueron estudiados con el modelo de Rasch. Los resultados indicaron que sus elementos se concentran en un nivel inferior de dificultad, corto de ítems difíciles de evaluar niveles superiores de percepción emocional. Tanto RMET como PEP mostraron correlaciones significativas con RA y RV, corroborando otros estudios que muestran que la percepción emocional está relacionada con otros tipos de inteligencia. Sin embargo, la correlación entre RMET y PEP fue menor que lo esperado ($r = .43$), lo que sugiere que percepción de emociones en los ojos está sólo parcialmente relacionada con percepción en toda la cara.

Palabras clave: expresiones faciales, inteligencia, evaluación psicológica, teoría de respuesta al ítem

Emotional expression is considered an important part of social interaction, with overt emotional behaviour functioning as nonverbal means of communication. The ability to express emotions conveys information regarding one's feelings, needs, intentions, and preoccupations, with specific bodily changes associated with their corresponding emotional states (Ekman, 2003; Matsumoto, Keltner, Shiota, O'Sullivan, & Frank, 2008; Niedenthal, Krauth-Gruber, & Ric, 2006; Strongman, 2003). This includes typical facial and bodily alterations whenever someone is, for instance, feeling angry during an unfair situation, or afraid of being exposed to a dangerous situation.

As important as expression is the ability to adequately recognize manifestations of emotional states. Emotional perception is shown to correlate with social abilities (Chen, 2014), academic achievement (MacCann, Fogarty, Zeidner, & Roberts, 2011), cognitive abilities (Jesus Junior & Noronha,

[2]Correspondence address:
Departamento de Psicologia e Psicanálise, Centro de Ciências Biológicas, Universidade Estadual de Londrina Campus Universitário. Caixa Postal 10011, CEP 86057-970. Londrina-PR, Brazil. E-mail: fabiano@avalpsi.com.br

Available in www.scielo.br/paideia

2007; Miguel, Ogaki, Inaba, & Ribeiro, 2013), among others. On the other hand, deficits in recognizing emotions are present in disorders such as autism (Sucksmith, Allison, Baron-Cohen, Chakrabarti, & Hoekstra, 2013), Down syndrome (Virji-Babul, Watt, Nathoo, & Johnson, 2012), eating disorders (Joos, Cabrillac, Hartmann, Wirsching, & Zeeck, 2009), depression (Langenecker et al., 2005), schizophrenia (J. Addington & D. Addington, 1998), and many others.

Tests of emotional recognition are usually administered in research, and the Reading the Mind in the Eyes Test (RMET) is one of the most popular instruments, displaying photos of the eyes and surrounding region of people expressing emotions. The test has shown evidences of validity and fair to good psychometric characteristics in numerous countries and languages (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Chapman et al., 2006; Fernández-Abascal, Cabello, Fernández-Berrocal, & Baron-Cohen, 2013; Irani et al., 2006; Pfaltz et al., 2013; Prevost et al., 2013; Vellante et al., 2013).

The RMET is generally considered a test of the theory of mind, which means the test taker must attribute a feeling to the people he/she sees in the photos based on the emotional expression of their eyes. Peterson and Miller (2012) suggested that such ability may not be completely unrelated to other cognitive abilities after finding a large magnitude correlation ($r = .49$) between the RMET and a verbal IQ task. Such results draw attention to a broader field of study, which is emotional intelligence. Emotional intelligence is considered the ability to adequately recognize emotions, to understand how they blend and change over time, and to regulate them in order to achieve well-being and better social interactions (Caruso, Salovey, Brackett, & Mayer, 2015; Mayer, Salovey, & Caruso, 2008). Thus, emotional perception is part of the emotional intelligence construct and considered an important ability that helps to promote adaptation.

Recent studies have shown that emotional intelligence, measured by performance tasks, correlate with other intelligence measures (Mayer, Roberts, & Barsade, 2008; Miguel, Ogaki et al., 2013), suggesting that it is a cognitive ability. In fact, analyses of structural models showed that emotional intelligence can be considered a second-stratum factor in the Cattell-Horn-Carroll model of intelligence, related to but distinct from other cognitive abilities such as fluid reasoning, comprehension-knowledge, visual processing, etc. (MacCann, Joseph, Newman, & Roberts, 2014).

In Brazil, only one research was found for the RMET. Sanvicente-Vieira et al. (2014) published a paper reporting the translation procedures and a pilot test of a computerized version with 10 people. The results showed that the scores of the computerized version mirrored the original paper-and-pencil version from international studies. However, the test is still lacking validity studies.

The goal of the present research was to study the psychometric properties of the RMET in a large Brazilian sample, with focus on item functioning. For that, the items were analysed using Rasch model, reporting item difficulties, fit, and differential item functioning. In addition, we studied the construct validity of the test by analysing its association with another emotional perception task and two tests of verbal and abstract reasoning. Taking into account findings in the field of emotional intelligence, we expected emotional perception ability to be correlated with other cognitive abilities.

Method

Participants

The number of participants was 1440, with ages ranging from 17 to 59 ($M = 21.8$, $SD = 6.6$), and 74.8% females. Regarding education level, 1.4% were students of elementary school, 30.4% were enrolled in or completed high school, 56.6% were enrolled in or completed college, and 11.7% were enrolled in or completed graduate education. Several courses were reported by the undergraduate and graduate students, with Psychology being the most prominent (43.6%), followed by Law (4.8%), Business Management (4.3%) and Medicine (3.2%). All other 40 courses were represented by less than 3% each.

Instruments

Reading the Mind in the Eyes Test (RMET). The RMET is composed of 37 pictures of the region of the eyes – including eyebrows and part of the nose – of people expressing emotions, with the first picture used as an example of the task. The participant must choose out of four words the one that describes how the person in the photo is feeling. Autism Research Centre's webpage (www.autismresearchcentre.com) provides a list of several languages that the RMET has been translated to, and states that the download is free as long as it is used for research purposes. According to their website, the Brazilian-portuguese version was translated by Tonelli and Sanvicente-Vieira. Nevertheless, we conducted a pilot study with four people regarding text comprehension, showing the translated version along with the original English version. These people had moderate to advanced knowledge of English, and their feedback was that they believed some of the alternatives carried different meanings in the translation, and that they would choose different alternatives on the test because of that. Thus, we consulted an English teacher and changed the labels of a few items because we considered their translations were either incorrect or unusual. They were: figure 5, worried was "atormentado", changed to "preocupado"; figure 12 and 33, embarrassed was "embaraçado", changed to "envergonhado"; figure 15, flustered was "cuidadoso", changed to "agitado"; figure 16, sympathetic was "compadecido", changed to "solidário"; figure 24, pensive was "apreensivo", changed to "pensativo"; figure 29, reflective was "absorto", changed to "refletindo"; figure 34, baffled was "perturbado", changed to "perplexo".

Computerized Test of Primary Emotions Perception (PEP). The test consists of 38 videos of people expressing emotions, displaying the head and upper part of the torso, with the first three videos used as examples. The videos last from 3 to 8 seconds. After watching each video, the participant must select the expressed emotions, from a total of eight: joy,

love, fear, surprise, sadness, disgust, anger, and curiosity. The instructions explain that the videos may contain more than one emotional expression.

Previous studies with the PEP have found significant correlations with abstract (*r* = .36) and verbal (*r* = .38) reasoning, and null correlations with self-report measures of personality (Miguel, Finoto, & Miras, 2013; Miguel, Ogaki et al., 2013). For the present sample, Rasch reliability was .70.

Abstract Reasoning (AR) and Verbal Reasoning (VR). AR and VR are parts of the Battery of Reasoning Tests, a set of cognitive tasks that has shown adequate evidences of validity (Flores-Mendoza et al., 2013; Primi, Couto, Almeida, Guisande, & Miguel, 2012; Primi, Silva, Rodrigues, Muniz, & Almeida, 2013). In each of the 28 items of AR, the participant is presented with two abstract geometric figures where the first one was transformed into the second one (for example, size changed, lines were added, etc.). The same transformation rule must be applied to a third picture in order to find the correct answer from five alternatives. VR displays 28 items composed of words that share a relationship (for example, "carpet is to floor the same way painting is to...?") and the participant must also choose from five alternatives. In both tests, the first three items serve as examples. The Rasch reliability was .74 for AR, and .64 for VR.

Procedure

Data collection. RMET, AR and VR are paper-and-pencil tests, and they were adapted to online format for this research. The online versions featured all the instructions from the original formats, and displayed only one item at a time. The PEP was originally created as an online test, so no change was made to its format.

Participants were recruited through Facebook and direct contact of undergraduate and graduate students. Following guidelines from International Test Commission (2005), participants registered on the system using their emails as login, and signed the online version of the agreement consent. After the agreement, an initial page displayed links to answer the four tests, and participants could choose which one they wanted to answer first. The research was online for nine months.

For all tests, participants would not be able to move to the next item if they left an item unanswered, although they were allowed to abandon the testing. For purpose of this research, only tests that were answered completely were studied. However, because of the nature of the system, not all participants answered all tests. We selected those that answered both the RMET and PEP, which were 1440 participants. From those, 844 answered AR, 827 answered VR, and 661 answered both reasoning tasks.

Data analysis. The answers to all tests were analyzed based on Item Response Theory, using Rasch model with software Winsteps. This model estimates the difficulty for each item of the test, and returns the participants' scores in the same scale. By default, the mean difficulty of the items of a test is .00. There are no standard terms to describe specific difficulty values, though the more they distance from the mean the more they are considered difficult (if positive) or easy (if negative).

For the RMET, we also estimated indices of misfit, point-measure correlation and differential item functioning (DIF). Misfit indicates when an item is behaving in an unproductive way, generally being removed from analyses. Linacre (2009) suggests that items with indices above 1.50 should be used with care, while indices above 2.00 should be considered degrading to the measure. Point-measure correlations are equivalent to item-total, indicating the amount of contribution of an item to the total measure. DIF is a comparison between groups to verify whether an item has too discrepant difficulties among them, which generally indicates that the item may privilege or hinder one or more of the groups. For this research, we ran DIF analyses comparing levels of schooling, gender, and format of test. For schooling, participants were divided in two groups: those with education until high school, and those with college education or higher. For test format, 30 participants answered the paper-and-pencil version of the RMET instead of the online version.

In regard to the other measures, we performed Pearson correlations to study convergent validity of the RMET. In order to maintain methodological coherence, the other instruments were also scored based on Rasch model. In addition, because the age range was broad, we conducted partial correlations controlling for age and educational level, as intelligence is known to be related to those data (Primi et al., 2012).

Ethical Considerations

The research was accepted by the Universidade Estadual de Londrina's Ethics Committee under protocol n. 191/2010 (CAAE n. 0159.0.268.000-10). The procedures were conducted only after its acceptance. Participants were shown the informed consent term on the computer screen. The term informed the procedure and confidentiality of the data, and that the results would be treated anonymously, with the possibility to abandon research at any time. After agreeing with participating, the tests were displayed.

Results

The items parameters of the RMET were calculated based on all 1440 participants that answered the test. The mean difficulty was .00 with a standard deviation of .81, and a maximum difficulty of 1.39 and minimum of -1.71. The maximum infit and outfit were 1.06 and 1.07, respectively. The Rasch reliability was .49, a lower value than the Cronbach's alphas found in previous international studies, although those indexes were not too high either, ranging from .58 to .70 (for a review see Vellante et al., 2013). Linacre (1997) states that tests such as Cronbach's alpha and KR-20 overestimate the true reliability, so a lower index in Rasch is expected. Still, considering that no item displayed misfit and the point-measure correlations were equal to or higher than .15, the low reliability was unexpected. We then removed items 1, 17, 23, 29, 35, and 36 – that displayed the lower point-measure correlations –, which resulted in a 30-item test. A

new analysis showed that the reliability raised to .52, and that was the highest reliability index found when experimenting removing other items. Thus, we used the 30-item version in the following analysis, and its psychometric properties are presented in Table 1.

Table 1
Rasch Measures of the Items of the Reading the Mind in the Eyes Test

Item	Percentage of correct	Difficulty	Std. error	Infit	Outfit	Pt-meas. corr.	DIF significance		
							By gender	By schooling	By test format
1	37.1								
2	42.9	1.22	.07	1.06	1.08	.19	.790	.643	.181
3	77.7	-0.44	.08	0.98	0.94	.27	.426	.848	.402
4	66.4	0.17	.08	1.02	1.02	.22	.740	.749	.988
5	75.0	-0.28	.08	1.02	1.05	.19	.086	.316	.762
6	78.6	-0.49	.09	0.98	0.96	.25	.809	.584	.391
7	51.8	0.83	.07	1.02	1.03	.25	.587	.382	.456
8	80.4	-0.61	.09	1.00	1.00	.21	.842	.137	.611
9	56.5	0.62	.07	1.02	1.02	.25	.623	.116	.955
10	51.0	0.86	.07	1.05	1.05	.21	.779	.101	.433
11	77.0	-0.39	.08	1.01	1.01	.21	.701	.470	.517
12	44.1	1.17	.07	1.04	1.05	.22	.409	.832	.023
13	77.0	-0.39	.08	1.00	1.00	.22	.707	.133	.299
14	75.0	-0.28	.08	0.98	0.95	.27	.690	.233	.474
15	91.9	-1.66	.13	0.99	0.88	.19	.567	.993	.691
16	60.3	0.45	.07	1.02	1.02	.24	.395	.740	.419
17	48.8								
18	87.2	-1.13	.10	1.00	0.94	.19	.699	.899	.082
19	59.2	0.50	.07	1.02	1.01	.24	.863	.185	.857
20	77.0	-0.39	.08	0.95	0.90	.32	.786	.675	.470
21	77.3	-0.41	.08	0.98	0.96	.26	.442	.169	.431
22	88.7	-1.28	.11	0.98	0.93	.21	.474	.421	.226
23	64.2								
24	68.1	0.09	.08	0.95	0.93	.33	.678	.472	.190
25	68.3	0.08	.08	1.03	1.05	.19	.724	.175	.151
26	62.8	0.34	.07	0.92	0.90	.38	.906	.343	.923
27	43.2	1.20	.07	0.99	0.99	.30	.612	.366	.897
28	72.0	-0.11	.08	1.00	1.01	.25	.327	.226	.536
29	74.2								
30	90.5	-1.48	.12	0.98	1.03	.18	.799	.195	.835
31	62.5	0.35	.07	1.04	1.04	.21	.435	.997	.019
32	56.4	0.63	.07	1.01	1.03	.25	.147	.294	.279
33	46.3	1.07	.07	1.01	1.02	.26	.545	.067	.406
34	74.0	-0.22	.08	0.94	0.90	.34	.962	.186	.456
35	42.3								
36	89.0								

The items were then studied for differential functioning (DIF), using as criteria gender, educational level and test format. There was no significant DIF for any items considering gender or schooling. Regarding test format, two items displayed significant DIF: items 12 was more difficult in web (1.20) than paper version (-.81), while item 31 was more difficult in paper (2.35) than web version (.35).

After the items were analyzed, participants' scores were calculated for all tests, and are presented in Table 2. The fact that the scores were all above the mean difficulty of items (.00) for RMET, PEP, AR, and VR show that, in general, participants tended to get most of the answers right than wrong.

Pearson and partial correlations between the instruments were conducted. The results are displayed in Table 3.

Table 2
Descriptive Statistics of Participants' Scores

	RMET	PEP	AR	VR
M	0.90	0.52	1.93	1.37
SD	0.50	0.50	1.23	0.88
Minimum	-0.61	-1.15	-3.46	-1.29
Maximum	2.88	2.49	5.50	4.03

Note. RMET = Reading the Mind in the Eyes Test, PEP = Computerized Test of Primary Emotions Perception, AR = Abstract Reasoning, VR = Verbal Reasoning.

Table 3
Pearson and Partial Correlations Between the Instruments

		General			Controlling for age			Controlling for education		
		RMET	PEP	AR	RMET	PEP	AR	RMET	PEP	AR
PEP	*r*	.43***			.41***			.42***		
	N	1440			1440			1440		
AR	*r*	.14***	.24***		.04	.15***		-.02	.21***	
	N	844	844		844	844		844	844	
VR	*r*	.19***	.23***	.45***	.15***	.19***	.43***	.07	.24***	.40***
	N	827	827	661	827	827	661	827	827	661

Note. RMET = Reading the Mind in the Eyes Test, PEP = Computerized Test of Primary Emotions Perception, AR = Abstract Reasoning, VR = Verbal Reasoning.
***$p < .001$.

Discussion

A closer inspection of the difficulties of RMET items showed discrepancies from the studies conducted in USA, Spain, Germany, France, and Italy (Baron-Cohen et al., 2001; Fernández-Abascal et al., 2013; Pfaltz et al., 2013; Prevost et al., 2013; Vellante et al., 2013). For example, before being removed, item 01 was the most difficult item of the test, with only 37.1% of the participants choosing the correct answer. Although it was not the easiest item in the international studies either, the percentage of correct answers ranged from 57.1% to 84%. On the other hand, item 11 was chosen correctly by 77.0% of the participants of the present research, while the international results showed a range from 57% to 74.3%. Thus, in general, the RMET displayed a pattern of item functioning that is equivalent to the international studies, i.e., a few items had similar percentages of correct answers, while other items had lower or higher percentages, which may reflect particularities of each country's culture. We hypothesize that this situation is created by the social understanding of the meaning of

words in each country, and possibly reflects the importance of specific emotion labels throughout cultures, as has been shown in other studies (Nelson & Russell, 2013). In this sense, even if words for emotions are correctly translated, different cultures may value specific expressions and have a tendency to recognize them better than other cultures.

Six items (1, 17, 23, 29, 35, 36) were removed from the RMET due to poor psychometric performance. It is noticeable that those items display words that are unique to them. For example, "animado" appears only in items 1 and 17; "nervoso" only in items 35 and 36; "desafiador" only in item 23; and so forth. A possible explanation is that such words are not frequently known or understood by the participants in daily life, which indicates low ecological validity. A further study of the RMET is then suggested, using different terms for those items and verifying whether their psychometrical properties improve. If there is no improvement, the emotions are probably poorly represented and the items are contributing to measurement errors, and so we recommend that further studies with the RMET use such items with attention.

DIF was found for two items, comparing test format.

Because one item was more difficult in paper version and the other one was more difficult in web version, it is possible that their effects cancel each other out when measuring the participant's ability. In addition, the number of test items that displayed DIF is low, which could hardly account for the test's low reliability.

The way that the items are distributed is a possible explanation for the low reliability. The participants' mean score in the RMET was .90, which is above 1 standard deviation of items difficulty (.81) and close to the item with the highest level of difficulty (item 2 = 1.22). Such results show that the RMET lacks items at a higher level of difficulty, leaving a little more than 50% of participants with few items to analyze their abilities, and even no items for people with abilities higher than 1.22, which is around 20% of the present sample. Because less items to evaluate an ability leads to larger measurement error, probably this is what is accounting for the test's low reliability in the current sample. Consistently, PEP, AR, and VR have larger standard deviations of items difficulty, thus properly covering a wide range of participants' abilities and displaying better reliability indices.

It is clear that the RMET needs items that are more difficult. A concentration of low difficulty items may still result in a good measure if one is assessing low levels of emotional perception, such as the ones that characterize autism or other disorders. However, the lack of difficult items hinders an accurate assessment of people with high levels of emotional perception, as was the case of this research's participants, resulting in a low reliability measure.

When compared with the other tests, the RMET displayed a moderate correlation with the PEP ($r = .43$), which indicates that the constructs measured by RMET and PEP are related, but are not exactly the same. Because both tests are supposed to assess the ability to identify emotions in other people, a slightly higher correlation was expected (above .50) (Cohen, 1992). Although the result still indicates a large shared variance between the instruments, it also suggests that different forms of assessing emotional perception are not highly correlated. In fact, if both instruments were highly correlated (for example, above .80), it would indicate that one is a parallel form of the other, not justifying the existence of two tests. So, our results denotes that emotional perception is affected by the nature of the stimuli, which in this research may be represented by the differing test formats – one displays pictures while the other displays videos.

Moreover, the RMET focuses on a particular region of the face in order to measure emotion recognition, while the PEP displays the whole face. It is known that culture may influence on the strategy to recognize emotions. For example, a study by Blais, Jack, Scheepers, Fiset, and Caldara (2008) with Western Caucasians and East Asians demonstrated that individuals from different cultures tend to focus on different regions of the face in order to identify the emotion being expressed. Accordingly, Beaudry, Roy-Charland, Perron, Cormier, and Tapp (2014) discovered that, depending on the expressed emotion, eyes and mouth contribute in different levels of importance for recognition. It may be the case that Brazilians use information from other regions besides the eyes. A study that tracked eyes movement during the task of emotion recognition would be required to test this assertion. If confirmed, it would mean that the RMET is not a reliable instrument for Brazilian culture, and assessment of emotional perception would require tests with larger or different regions of the face, such as the PEP.

A correlation between RMET and VR was expected. Our results replicate the findings by Peterson and Miller (2012), and indicates that a better understanding of words is mildly associated with correct naming of emotions. The RMET was also associated with AR, albeit in a lower level than with VR. These results are in accordance with those found in literature, specially related to emotional intelligence, in which emotional perception is correlated with other cognitive abilities. Slightly higher correlations were also found between the PEP and both AR and VR, which indicates that the ability to recognize expressions of emotions shares some variance with other cognitive abilities.

Considering the partial correlations, the association between RMET and PEP were similar when controlling for both age and education, as well as both cognitive tasks (AR and VR). These results were expected, as the two dyads (RMET x PEP, and AR x VR) involve similar tasks. However, the association between RMET and AR and VR were different. All magnitudes were non-significant, with the exception of RMET and VR controlling for age, that nonetheless was lower. These results imply that there is influence of both age and educational factors in the association of RMET with cognitive tasks. While age and education impact is expected for intelligence tests (Primi et al. 2012), the null correlations for RMET suggest that the construct assessed by the RMET is highly influenced by age and education, which may represent the cognitive development.

On the other hand, the partial correlations between the PEP and the two cognitive tasks remained significant, although some were lower, indicating that emotional perception measured by the PEP still shares variance with intelligence, even when controlling for age and education. Such results are in accordance with those found by MacCann et al. (2014). The authors analysed several structural models that included measures of emotional intelligence and other cognitive tasks such as fluid and crystalized intelligence, quantitative and visual reasoning, among others, and confirmed a second-stratum factor for emotional intelligence. Although our research did not administer broader measures of emotional intelligence, the perception of emotions is considered part of the construct (Caruso et al., 2015; Mayer, Salovey et al., 2008). The shared variance gives support to the suggestion that emotional perception in facial expressions is part of a particular set of cognitive abilities, which is emotional intelligence. The findings from this study have implications for cognitive assessment, such as the need to include emotional perception tasks in a battery of tests. In addition, considering that identification of emotions is part of social interactions, tests of emotional perception should be included in assessment of social adaptation skills or problems, as corroborated by previous studies (J. Addington & D. Addington, 1998; Chen, 2014; Joos et al., 2009;

Langenecker et al., 2005; MacCann et al., 2011; Sucksmith et al., 2013; Virji-Babul et al., 2012). However, as discussed above, while the PEP seems to be a cognitive measure of emotional intelligence, the RMET seems not to be. In this sense, it could be considered a test of theory of mind as it was originally designed, which is not considered a reasoning ability, but the process of putting oneself in another's shoes and adjusting to their mental states (Baron-Cohen et al., 2001; Fernández-Abascal et al., 2013).

One possible limitation of the research is that both measures of emotional perception rely on written instructions, which require vocabulary knowledge. This was particular true for the RMET, which showed lower correlations with intelligence after controlling for age and education. Thus, the covariance with verbal reasoning may indicate that this ability is required to answer the test, but may not be related to pure perception of emotions. A way to test this hypothesis is to develop tests that do not require reading comprehension.

Another limitation is that all participants had internet access to the tests, with the exception of 30 people that answered the paper-and-pencil version of the RMET. While DIF analyses showed that both versions are not significantly different, we did not assess individuals that do not have internet access, such as inpatients, poor or uneducated people. Further studies are required to broaden the tests' applicability, and to understand their relations to other constructs, such as personality.

References

Addington, J., & Addington, D. (1998). Facial affect recognition and information processing in schizophrenia and bipolar disorder. *Schizophrenia Research, 32*(3), 171-181. doi:10.1016/S0920-9964(98)00042-5

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241-251. doi:10.1111/1469-7610.00715

Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., & Tapp, R. (2014). Featural processing in recognition of emotional facial expressions. *Cognition and Emotion, 28*(3), 416-432. doi:10.1080/02699931.2013.833500

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS One, 3*(8), e3022. doi:10.1371/journal.pone.0003022

Caruso, D. R., Salovey, P., Brackett, M., & Mayer, J. D. (2015). The ability model of emotional intelligence. In S. Joseph (Ed.), *Positive psychology in practice: Promoting human flourishing in work, health, education, and everyday life* (2nd ed., pp. 545-558). Hoboken, NJ: John Wiley & Sons.

Chapman, E., Baron-Cohen, S., Auyeung, B., Knickmeyer, R., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy: Evidence from the Empathy Quotient (EQ) and the "Reading the Mind in the Eyes" test. *Social Neuroscience, 1*(2), 135-148. doi:10.1080/17470910600992239

Chen, J. (2014). Face recognition as a predictor of social

cognitive ability: Effects of emotion and race on face processing. *Asian Journal of Social Psychology, 17*(1), 61-69. doi:10.1111/ajsp.12041

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. doi:10.1037/0033-2909.112.1.155

Ekman, P. (2003). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life.* New York, NY: Times Books.

Fernández-Abascal, E. G., Cabello, R., Fernández-Berrocal, P., & Baron-Cohen, S. (2013). Test-retest reliability of the "Reading the Mind in the Eyes" test: A one-year follow-up study. *Molecular Autism, 4*(33). doi:10.1186/2040-2392-4-33

Flores-Mendoza, C., Widaman, K. F., Rindermann, H., Primi, R., Mansur-Alves, M., & Pena, C. C. (2013). Cognitive sex differences in reasoning tasks: Evidence from Brazilian samples of educational settings. *Intelligence, 41*(1), 70-84. doi:10.1016/j.intell.2012.11.002

Irani, F., Platek, S. M., Panyavin, I. S., Calkins, M. E., Kohler, C., Siegel, S. J., . . . Gur, R. C. (2006). Self-face recognition and theory of mind in patients with schizophrenia and first-degree relatives. *Schizophrenia Research, 88*(1-3), 151-160. doi:10.1016/j.schres.2006.07.016

International Test Commission. (2005). *ITC guidelines on computer-based and internet delivered testing*. Retrieved from https://www.intestcom.org/files/guideline_computer_based_testing.pdf

Jesus Junior, A. G., & Noronha, A. P. P. (2007). Inteligência emocional e provas de raciocínio: Um estudo correlacional [Emotional intelligence and reasoning tests: A correlacional study]. *Psicologia: Reflexão e Crítica, 20*(3), 480-489. doi:10.1590/S0102-79722007000300016

Joos, A. A. B., Cabrillac, E., Hartmann, A., Wirsching, M., & Zeeck, A. (2009). Emotional perception in eating disorders. *International Journal of Eating Disorders, 42*(4), 318-325. doi:10.1002/eat.20621

Langenecker, S. A., Bieliauskas, L. A., Rapport, L. J., Zubieta, J.-K., Wilde, E. A., & Berent, S. (2005). Face emotion perception and executive functioning deficits in depression. *Journal of Clinical and Experimental Neuropsychology, 27*(3), 320-333. doi:10.1080/13803390490490515720

Linacre, J. M. (1997). KR-20 / Cronbach alpha or Rasch person reliability: Which tells the "truth"? *Rasch Measurement Transactions, 11*(3), 580-581.

Linacre, J. M. (2009). *A user's guide to Winsteps-Ministep: Rasch-model computer programs*. Chicago, IL: Winsteps.com.

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*(1), 60-70. doi:10.1016/j.cedpsych.2010.11.002

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. doi:10.1037/a0034755

Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M., & Frank, M. (2008). Facial expressions of emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 211-234). New York, NY: Guilford.

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

Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence: New ability or eclectic traits? *American Psychologist, 63*(6), 503-517. doi:10.1037/0003-066X.63.6.503

Miguel, F. K., Finoto, B. A. S., & Miras, B. D. (2013). Percepção emocional e traços de personalidade: Estudo de validade divergente [Emotional perception and personality traits: Divergent validity study]. *Encontro: Revista de Psicologia, 16*(24), 107-120.

Miguel, F. K., Ogaki, H. A., Inaba, C. M., & Ribeiro, D. O. (2013). Percepção emocional e inteligência: Contribuições para o modelo CHC [Emotional perception and intelligence: Contributions to the CHC model]. *Revista Sul-Americana de Psicologia, 1*(1), 36-47.

Nelson, N. L., & Russell, J. A. (2013). Universality revisited. *Emotion Review, 5*(1), 8-15. doi:10.1177/1754073912457227

Niedenthal, P. M., Krauth-Gruber, S., & Ric, F. (2006). *Psychology of emotion: Interpersonal, experiential, and cognitive approaches*. New York, NY: Psychology Press.

Peterson, E., & Miller, S. (2012). The Eyes Test as a measure of individual differences: How much of the variance reflects verbal IQ? *Frontiers in Psychology, 3,* 220. doi:10.3389/fpsyg.2012.00220

Pfaltz, M. C., McAleese, S., Saladin, A., Meyer, A. H., Stoecklin, M., Opwis, K., . . . Martin-Soelch, C. (2013). The Reading the Mind in the Eyes Test: Test-retest reliability and preliminary psychometric properties of the German version. *International Journal of Advances in Psychology, 2*(1), 1-9.

Prevost, M., Carrier, M.-E., Chowne, G., Zelkowitz, P., Joseph, L., & Gold, I. (2013). The Reading the Mind in the Eyes test: Validation of a French version and exploration of cultural variations in a multi-ethnic city. *Cognitive Neuropsychiatry, 19*(3), 189-204. doi:10.1080/13546805.2013.823859

Primi, R., Couto, G., Almeida, L. S., Guisande, M. A., & Miguel, F. K. (2012). Intelligence, age and schooling: Data from the Battery of Reasoning Tests (BRT-5). *Psicologia: Reflexão e Crítica, 25*(1), 79-88. doi:10.1590/S0102-79722012000100010

Primi, R., Silva, M. C. R., Rodrigues, P., Muniz, M., & Almeida, L. S. (2013). The use of the bi-factor model to test the uni-dimensionality of a battery of reasoning tests. *Psicothema, 25*(1), 115-122. doi:10.7334/psicothema2011.393

Sanvicente-Vieira, B., Kluwe-Schiavon, B., Wearick-Silva, L. E., Piccoli, G. L., Scherer, L., Tonelli, H. A., & Grassi-Oliveira, R.

(2014). Revised Reading the Mind in the Eyes Test (RMET): Brazilian version. *Revista Brasileira de Psiquiatria, 36*(1), 60-67. doi:10.1590/1516-4446-2013-1162

Strongman, K. T. (2003). *The psychology of emotion: From everyday life to theory*. Chichester, United Kingdom: John Wiley & Sons.

Sucksmith, E., Allison, C., Baron-Cohen, S., Chakrabarti, B., & Hoekstra, R. A. (2013). Empathy and emotion recognition in people with autism, first-degree relatives, and controls. *Neuropsychologia, 51*(1), 98-105. doi:10.1016/j.neuropsychologia.2012.11.013

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., & Preti, A. (2013). The "Reading the Mind in the Eyes" test: Systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry, 18*(4), 326-354. doi:10.1080/13546805.2012.721728

Virji-Babul, N., Watt, K., Nathoo, F., & Johnson, P. (2012). Recognition of facial expressions of emotion in adults with Down Syndrome. *Physical & Occupational Therapy in Pediatrics, 32*(3), 333-343. doi:10.3109/01942638.2011.653626

Fabiano Koich Miguel is an Assistant Professor of the Universidade Estadual de Londrina.

Raissa Barquete Caramanico is an undergraduate student of the Universidade Estadual de Londrina.

Eduardo Yudi Huss is an undergraduate student of the Universidade Estadual de Londrina

Ana Carolina Zuanazzi is a Ph.D. candidate of the Graduate Program in Psychology at the Universidade São Francisco.