

Comparação entre dois sistemas de pontuação para o teste informatizado de percepção de emoções em fotos

Comparison between two scoring systems for the computerized emotional perception test using photographs

José Maurício Haas **BUENO**¹

Fabiano Koich **MIGUEL**²

Ricardo **PRIMI**²

Monalisa **MUNIZ**¹

Gleiber **COUTO**²

Ana Paula Porto **NORONHA**²

Resumo

Este trabalho teve o objetivo principal de comparar dois sistemas de atribuição de pontos às respostas de sujeitos no Teste Informatizado de Percepção de Emoções: concordância com o consenso e concordância com critérios definidos por especialistas, com base na teoria relacionada à inteligência emocional e na teoria de resposta ao item. Participaram do estudo 55 sujeitos de ambos os sexos, de 18 a 74 anos, cujos protocolos foram avaliados pelos dois métodos de pontuação. Embora a correlação entre os dois métodos de pontuação tenha sido elevada, a atribuição de pontos por concordância com especialistas resultou em correlações item-total significativamente mais elevadas que as pontuações por concordância com o consenso. Os resultados sugerem que o desenvolvimento de instrumentos para avaliação de habilidades relacionadas à inteligência emocional deve priorizar o critério da concordância com especialistas, e mostra como a teoria de resposta ao item pode ser útil no processo de investigação e escolha desses critérios.

Unitermos: Inteligência emocional. Precisão do teste. Teoria de resposta ao item.

Abstract

The main aim of this study was to compare two rating systems relating to the responses of individuals in the Computerized Emotional Perception Test using Photographs: consensus agreement and agreement with criteria defined by experts based on emotional intelligence theory and Item Response Theory. Participants comprised 55 subjects of both sexes, aged between 18 and 74, whose protocols were rated by both rating methods. Although a high correlation between the two rating methods was observed, the expert agreement criteria yielded significantly higher item-total correlations than consensus agreement. The results suggest that the development of evaluation tools for emotional intelligence related skills should be based primarily on expert agreement criteria, and shows how the Item Response Theory should be helpful in this process of investigation and selection of criteria.

Uniterms: Emotional intelligence. Test reliability. Item response theory.

▼▼▼▼▼

¹ Universidade São Francisco, Curso de Psicologia, Laboratório de Avaliação Psicológica e Educacional. R. Alexander Rodrigues Barbosa, 45, Campus Itatiba, Centro, 13251-900, Itatiba, SP, Brasil. Correspondência para/Correspondence to: J. M. H. BUENO. E-mail: <talktomav@gmail.com>.

² Universidade São Francisco, Programa de Pós-Graduação *Stricto Sensu* em Psicologia. Itatiba, SP, Brasil.

Desde sua proposição por Salovey e Mayer (1990), o conceito de inteligência emocional já passou por questionamentos e revisões que têm contribuído para seu entendimento na psicologia como uma nova forma de inteligência. Nesse sentido, a inteligência emocional é entendida como um conjunto de habilidades intercorrelacionadas e organizadas em níveis hierárquicos, de acordo com a complexidade dos processos psicológicos envolvidos. No primeiro nível estão as habilidades relacionadas à percepção, avaliação e expressão da emoção (percepção de emoções); no segundo nível, as habilidades relacionadas à utilização da emoção como facilitadora do pensamento (facilitação do pensamento); no terceiro nível está a habilidade de compreender e analisar as emoções, bem como empregar o conhecimento emocional (conhecimento emocional); por fim, tem-se a habilidade relacionada ao controle reflexivo de emoções para promover o crescimento emocional e intelectual (regulação emocional) (Mayer & Salovey, 1997).

Essas habilidades não são novidades na psicologia, exceto pela sua reunião em um campo maior de estudo sob o enfoque do processamento da informação, que, neste caso, é a emocional (Mayer & Salovey, 1993). A percepção de emoções, por exemplo, objeto deste estudo, tem sido foco de atenção de diversos pesquisadores, especialmente entre os que se preocupam com a capacidade de perceber emoções em expressões faciais (Ekman, 1965).

Lane (2000) sugere que a percepção de emoções pode ser um componente particularmente importante, ou primário, da inteligência emocional, por ser a base sobre a qual os outros processos a ela relacionados se edificam. A falha em detectar, avaliar e nomear o estímulo carregado de afeto implica em falha geral no processamento subsequente da informação emocional.

Dentro do escopo de estudo da inteligência emocional, o nível da percepção, avaliação e expressão da emoção abrange desde a capacidade de identificar emoções em si mesmo, em outras pessoas e em objetos ou condições físicas, até a capacidade de expressar essas emoções e as necessidades a elas relacionadas, incluindo a capacidade de avaliar a autenticidade de uma expressão emocional, detectar sua veracidade, falsidade ou tentativa de manipulação (Mayer & Salovey, 1997).

Há estudos mostrando que a percepção de emoções é uma habilidade unifatorial (Bueno & Primi, 2003; Davies, Stankov & Roberts, 1998; Mayer, DiPaolo & Salovey, 1990; Primi, Bueno & Muniz, 2006) e pelo menos um estudo evidenciando que esse tipo de processamento envolve estruturas neurais distintas das envolvidas no processamento de informações espaciais (Lane, Fink, Chua & Dolan, 1997). Esses autores mostraram que o foco da atenção na própria experiência emocional está associado a uma maior atividade neural do córtex cingulado anterior cefálico e do córtex pré-frontal mediano, ao passo que o foco da atenção em informações espaciais associou-se à ativação bilateral do córtex parieto-occipital, região associada às avaliações espaciais.

Alguns instrumentos foram desenvolvidos para avaliação objetiva da capacidade de perceber emoções por meio do desempenho dos sujeitos em tarefas cuja resolução requeira o processamento perceptual de informações emocionais. Por exemplo, Mayer et al. (1990) desenvolveram um procedimento para avaliação da percepção de emoções em estímulos não verbais (cores, desenhos abstratos e expressões faciais). Esse instrumento era constituído por seis expressões faciais, seis cores (vermelho, azul, verde, amarelo, preto e branco) e seis desenhos abstratos em preto e branco, totalizando dezoito estímulos. Diante de cada estímulo, apresentado individualmente, o sujeito devia atribuir pontos (de 1 a 5) a cada uma das seis emoções primárias (alegria, tristeza, raiva, medo, desgosto e surpresa), conforme a considerasse definitivamente não presente (1) até definitivamente presente (5). Nessas tarefas, o sujeito tinha que demonstrar sua habilidade relacionada à percepção de emoções, escolhendo não só qual ou quais as emoções presentes em cada estímulo, como também indicando suas respectivas intensidades.

Em outro estudo, Mayer e Geher (1996) utilizaram um sistema semelhante para avaliar a percepção de emoções em um conjunto de estímulos verbais. Os participantes desse estudo leram relatos breves de situações enfrentadas pelas pessoas no dia-a-dia e tentaram estimar quais os sentimentos experimentados pelos protagonistas das histórias. Cada relato era seguido por uma escala de 12 itens bipolares, na qual os participantes assinalavam suas respostas de forma dicotômica: experienciada mais intensamente *versus*

experienciada com menos intensidade. Nessa escala, portanto, o sujeito tinha que demonstrar apenas sua habilidade em escolher corretamente as emoções presentes.

O primeiro instrumento baseado em desempenho para avaliação da inteligência emocional lançado comercialmente foi a *Multifactor Emotional Intelligence Scale* (MEIS) (Mayer, Salovey & Caruso, 1997). Esse instrumento era composto por 12 tarefas destinadas a investigar quatro ramificações da inteligência emocional (Mayer & Salovey, 1997): identificação das emoções, utilização das emoções, compreensão das emoções e gerenciamento das emoções. A primeira ramificação desse instrumento, identificação de emoções, era constituída por tarefas semelhantes aos dois primeiros estudos relatados anteriormente (Mayer et al., 1990; Mayer et al., 1997), envolvendo, portanto, a identificação de emoções em estímulos verbais e não verbais.

O *Mayer-Salovey-Caruso Emotional Intelligence Test* (MSCEIT, V 2.0) (Mayer, Salovey & Caruso, 2002) é uma versão aprimorada da MEIS, em que cada um dos quatro eixos da inteligência emocional anteriormente descritos é avaliado por dois tipos de tarefas, totalizando oito seções identificadas por letras, de A até H. A percepção de emoções é avaliada nas seções A e E, a utilização da emoção para facilitação do pensamento é avaliada nas seções B e F, a compreensão emocional nas seções C e G, e a regulação emocional, nas seções D e H.

Embora a percepção de emoções tenha sido definida como a habilidade para perceber emoções em si mesmo e em outras pessoas, objetos ou condições físicas, as duas seções do MSCEIT destinadas à avaliação desse construto são destinadas somente à avaliação de emoções em condições externas: outras pessoas (seção A) e condições físicas (seção E). A percepção de emoções em expressões faciais é avaliada na sessão A, por meio de uma tarefa que consiste em apresentar uma expressão facial acompanhada de cinco possíveis sentimentos expressos por ela (por exemplo, felicidade, medo, surpresa, aversão e excitação). O testando é solicitado a registrar, em uma escala *Likert* de cinco pontos, a intensidade com que cada sentimento está sendo expresso em cada face (por exemplo, 1 - nenhuma felicidade até 5 - extrema felicidade). São apresentadas quatro expressões faciais, nas quais o respondente deve avaliar cinco sentimentos, totalizando 20 itens. A per-

cepção de emoções em paisagens é avaliada da mesma forma, na sessão E, com a diferença de que o estímulo é composto por paisagens e pinturas abstratas, ao invés de expressões faciais. No total são apresentadas seis imagens, nas quais o respondente deve avaliar cinco sentimentos, totalizando 30 itens.

Por serem medidas diretas da acuidade da percepção emocional apresentada pelas pessoas, os instrumentos de avaliação por desempenho têm sido preferidos aos instrumentos de avaliação por autorrelato. No entanto, há uma dificuldade relacionada à seleção de um critério para atribuição de pontos, ou da resposta que será considerada como correta. Cinco métodos de atribuição de pontos foram encontrados na literatura (Mayer et al., 1990; Mayer & Geher, 1996). O mais utilizado em pesquisas é baseado na *concordância com o consenso*, segundo o qual a pontuação do sujeito em cada item é proporcional à porcentagem de sujeitos que escolheram a mesma opção que ele. Outra possibilidade é utilizar o método da *concordância com especialistas*, em que o sujeito recebe um ponto sempre que sua resposta concorda com a avaliação realizada por profissionais (especialistas), que utilizam pesquisas e teorias relacionadas ao tema para definir uma resposta correta. Semelhante ao anterior, o método da *concordância com a pessoa-alvo* também atribui um ponto ao sujeito toda vez que sua resposta concordar com a resposta da pessoa que produziu o estímulo utilizado na tarefa (por exemplo, um relato verbal, uma expressão facial). Outros dois critérios, pouco utilizados, são da *intensidade* e da *amplitude* das emoções percebidas, que correspondem à média e ao desvio-padrão das respostas dos sujeitos ao longo de todos os itens de cada escala, respectivamente.

A convergência entre os métodos de pontuação foi analisada em vários estudos. Em um deles, foram encontradas baixas correlações entre as pontuações por consenso, intensidade e amplitude (Mayer et al., 1990). Nesse estudo, obtiveram-se propriedades psicométricas mais adequadas com o método de concordância com o consenso. Por isso, esse método passou a ser empregado de forma mais sistemática pelos autores em suas pesquisas subsequentes, em detrimento dos outros dois tipos de pontuação. Em outros estudos, no entanto, os pesquisadores relatam ter encontrado convergência entre os sistemas de pontuação por concordância com

o consenso e com especialistas (Mayer et al., 2002; Mayer, Salovey, Caruso & Sitarenios, 2003), e entre concordância com o consenso e com a pessoa-alvo na avaliação da subescala histórias (Mayer & Geher, 1996). Em todos os casos, ainda, a pontuação baseada na concordância com o consenso apresentou os melhores resultados em relação à consistência interna.

Há algumas reflexões sobre esses critérios de pontuação, especialmente o da concordância com o consenso, que tem sido o mais utilizado (Bueno & Primi, 2003; MacCann, Roberts, Matthews & Zeidner, 2004). Apesar de esse sistema resultar em boas propriedades psicométricas, ele não permite afirmar que um sujeito tenha maior ou menor acuidade perceptual das emoções, mas que ele concorda em maior ou em menor grau com a maioria das pessoas quanto a uma determinada percepção. O critério da resposta consensual parece mais adequado para avaliação de traços de personalidade, como no Rorschach, por exemplo, cujas respostas frequentes são interpretadas como respostas dadas por sujeitos que compartilham a mesma observação da realidade com a maioria das pessoas e que, portanto, teriam maior facilidade para adaptação.

No caso da avaliação de um tipo de inteligência, nem sempre uma resposta compartilhada com a maioria da população é a melhor. Por exemplo, supondo-se que em um teste de inteligência a dificuldade vá aumentando de item para item, é esperado que apenas uma minoria acerte os itens mais difíceis, que, segundo o critério de concordância com o consenso, receberiam pontuações menores do que a maioria que erra tais itens. Essas reflexões levaram os autores a apontar que a pontuação por concordância com especialistas é um ponto que merece ser investigado e desenvolvido, para que os instrumentos de avaliação da inteligência emocional por desempenho aproximem-se de algo mais adequado para avaliação da inteligência. Por isso, os objetivos deste trabalho foram: a) definir respostas consideradas corretas para um teste de percepção de emoções, a partir da análise de especialistas, com auxílio da teoria de resposta ao item (TRI), da teoria psicoevolutiva das emoções (Plutchik, 2002) e da teoria sobre inteligência emocional (Mayer & Salovey, 1997); e b) comparar as propriedades psicométricas entre as pontuações obtidas por concordância com especialistas e as obtidas por concordância com o consenso.

Método

Participantes

O grupo de especialistas, que analisou as respostas e estabeleceu o gabarito do teste, era composto por dois professores doutores e quatro estudantes de doutorado de um programa de pós-graduação em psicologia de uma cidade do interior paulista. Entre os componentes, dois eram do sexo feminino, uma professora e uma estudante, e quatro do sexo masculino, um professor e três estudantes.

O instrumento foi aplicado em 55 pessoas de ambos os sexos; 47,3% eram do sexo masculino e 52,7% do sexo feminino, constituindo uma amostra de conveniência encontrada na cidade de São Paulo. A idade variou de 18 a 74 anos, com média igual a 29,3 e desvio-padrão igual a 14,2. A maioria dos participantes tinha nível superior de escolaridade (72,7%), embora 18,2% tenham apresentado nível médio e 3,6%, nível fundamental.

Instrumentos

Para a coleta de dados, foi utilizado o Teste Informatizado de Percepção de Emoções em Fotos (Miguel, Primi, Muniz, Couto & Noronha, 2006), composto por 93 imagens que foram apresentadas na tela de um programa de computador. O teste é dividido em duas partes: as primeiras 36 imagens são utilizadas para avaliar a percepção da emoção presente nas personagens da foto (percepção de emoção no outro) e as 57 subsequentes são apresentadas para que o participante indique que emoção a imagem causa em si (percepção de emoção em si). No início da primeira parte, aparecem as instruções informando que as fotos são apresentadas uma a uma, e que a tarefa do sujeito é observá-las e julgá-las quanto à emoção que expressam. Na parte inferior da tela, abaixo da imagem apresentada, o sujeito escolhe uma entre nove opções; oito são relacionadas às emoções básicas propostas por Plutchik (2002) (raiva, medo, alegria, surpresa, tristeza, aversão, aceitação e expectativa) e uma à opção "neutro", caso o sujeito considere que a foto não expressa emoção. Após escolher uma opção de resposta, o sujeito deve clicar na opção "próxima" para que o programa apresente a

imagem seguinte. Ao fim da primeira parte, uma nova instrução surge na tela solicitando ao sujeito que, a partir daquele momento, observe as imagens e assinale as emoções que elas despertam em si, podendo haver algumas imagens repetidas da primeira sessão. Em seguida, as 57 imagens da segunda sessão são apresentadas uma a uma. As respostas dos sujeitos são armazenadas automaticamente em um banco de dados.

Procedimentos

O instrumento foi aplicado por alunos de graduação do curso de psicologia de uma grande universidade particular da cidade de São Paulo, devidamente instruídos e treinados para essa tarefa. A amostra foi de conveniência, controlada apenas quanto ao gênero. Os sujeitos convidados foram esclarecidos quanto aos objetivos e procedimentos da pesquisa e os que concordaram em participar assinaram um Termo de Consentimento Livre e Esclarecido, consentindo a utilização de seus dados para fins de pesquisa. Os dados das aplicações, registrados automaticamente, foram transferidos para uma planilha eletrônica, a partir da qual foram realizadas análises estatística descritivas e inferencial, com emprego do método de Rasch.

Para cada participante, foram obtidas duas pontuações pelo critério de concordância com o consenso e cinco pontuações por concordância com especialistas. Pelo critério de concordância com o consenso, cada participante recebeu uma pontuação proporcional ao número de pessoas que fizeram a mesma opção que ele. Por exemplo, se um participante escolheu uma opção que também foi escolhida por 60% das pessoas, então seu score naquele item foi de 0,60. Cada participante recebeu duas pontuações totais por esse critério, uma referente à percepção das emoções presentes nos personagens das fotos (chamada de Ct_outra nesta pesquisa), e outra referente à percepção da emoção que a imagem da foto causava no próprio observador (Ct_em_si).

Também foi utilizado o método de concordância com especialistas, cuja pontuação para cada item foi dicotômica: um para acerto e zero para erro. Cada participante recebeu cinco pontuações totais por concordância com especialistas. As duas primeiras foram semelhantes às pontuações por consenso: uma rela-

cionada ao total de vezes em que suas percepções das emoções presentes nas fotos concordavam com as definidas pelos especialistas (Et_outra), e outra referente ao número de vezes em que suas percepções das emoções que as fotos causavam em si concordavam com as definidas pelos especialistas (Et_em_si). Outras duas pontuações atribuídas estavam relacionadas à incapacidade de perceber emoções. Uma delas ocorria quando o participante optava pelo *neutro* em estímulos nos quais nitidamente havia uma emoção presente. Nesses casos, o total reflete o número de vezes que o sujeito percebeu uma emoção neutra quando, na opinião dos especialistas, deveria perceber uma emoção qualquer na foto (EN_outra) ou em si mesmo (EN_em_si). Finalmente, uma última pontuação atribuída estava relacionada ao total de percepções dinâmicas (t_din), definidas como a distorção perceptiva causada pelo dinamismo emocional do sujeito que percebe (Cattell, 1967). Nesses casos, na opinião de especialistas, é muito provável que o participante tenha imposto ao personagem da imagem apresentada uma emoção que a imagem despertava em si.

Resultados e Discussão

A apresentação dos resultados foi dividida em dois blocos. O primeiro refere-se à atribuição de pontos por concordância com especialistas, em que são apresentados um exemplo de como a análise de Rasch foi utilizada para avaliação dos itens e a definição dos critérios de pontuação por concordância com especialistas. O segundo bloco refere-se à comparação entre esse sistema de atribuição de pontos e o de concordância com o consenso, em que são apresentados os coeficientes de correlação de Pearson entre as diversas formas de pontuação empregadas, as correlações item-total entre as pontuações por consenso e por especialistas, nas fotos e nas pessoas, e os resultados *t* das comparações estatísticas entre os coeficientes de correlação item-total.

A definição das respostas corretas foi realizada por seis especialistas que analisaram todos os itens quanto ao conteúdo e aos resultados da análise de Rasch. Quando havia conflito entre a definição de uma resposta correta pelo conteúdo e pelos indicadores da análise de Rasch, prevalecia o critério de adequação da

resposta às proposições teóricas da teoria evolutiva das emoções (Plutchik, 2002) e da inteligência emocional (Mayer & Salovey, 1997). Por esses critérios, foram definidas uma ou duas respostas aceitas como corretas para cada item.

O método de Rasch forneceu os parâmetros empíricos para a análise dos itens. O coeficiente de correlação item-total forneceu um indicador de quanto o item contribuía com uma avaliação relacionada ao construto desejado. A comparação entre a média de *theta* (média de habilidade dos sujeitos) e as opções de respostas forneceu indicadores para comparação da habilidade dos sujeitos (*theta*) com as opções e pontuações que recebiam em cada item; já o índice de *outfit* serviu para a identificação do ajuste dos itens, ou seja, itens fáceis respondidos incorretamente por sujeitos com alta habilidade ou itens difíceis respondidos corretamente por sujeitos com baixa habilidade.

Por meio desses resultados, foi observado que alguns itens apresentavam características impróprias, como baixas correlações item-total, altos valores de *outfit* e média de *theta* alta para opções de resposta incorretas. Os itens identificados por um desses parâmetros como inapropriados foram confrontados pelos pesquisadores com a fundamentação teórica sobre emoções. Essa análise resultou na eliminação de seis itens do subtteste de percepção de emoções na foto e quinze do subtteste de percepção de emoções em si. O procedimento de eliminação pode ser visualizado no exemplo do item 22 (Tabela 1).

Como pode ser observado, esse item poderia ter duas respostas corretas, quais sejam, medo e aversão, e estas obtiveram correlação item-total abaixo de 0,15, considerada inadequada. Outro dado que contribuiu para a eliminação do item foi a média de habilidade (*theta*) dos indivíduos necessária para acertar o item.

Pode-se observar que os indivíduos com maior habilidade (*theta*) em perceber emoções nos outros erravam esse item (escore zero). Ao observar as categorias de resposta ao item, as respostas erradas assinaladas pelos indivíduos de maior *theta* foram examinadas, de modo a verificar se havia coerência entre a resposta dada e a tarefa proposta. Quando as opções eram coerentes, o item era mantido; quando as opções eram incoerentes, eliminava-se o item. Considerando a informação dos parâmetros e a análise teórica, ponderou-se que o item deste exemplo não ficou adequado para compor o subtteste. Por isso, o conjunto dessas informações levou à eliminação do item. Para os outros itens removidos, também se seguiu o mesmo padrão de análise e eliminação.

Após a análise de todos os itens, o subtteste de percepção de emoção no outro ficou composto por 30 itens e o subtteste de percepção de emoção em si, por 42 itens. Ao lado disso, foram calculadas as pontuações de percepções neutras em si e nos outros e o escore de percepção dinaceptiva, para uma melhor exploração dos dados. Na Tabela 2 são apresentadas as propriedades psicométricas de todas as pontuações obtidas pelo método de atribuição de pontos por concordância com especialistas.

Os dados referentes ao subtteste de percepção de emoções no outro (Et_outro) mostram que o índice de precisão pelo Rasch foi de 0,59, e pelo Coeficiente de Kuder-Richardson, de 0,68, considerados baixos pelos critérios internacionais, embora possam ser utilizados para fins de pesquisas. O valor médio dos *thetas* dos participantes sugere que o teste, cuja dificuldade média é centrada em zero pelo modelo de Rasch, tem uma tendência a ser mais fácil para esses sujeitos. Também os valores de ajuste (*outfit*) encontram-se dentro da faixa de adequação (Wright & Stone, 2004).

Tabela 1. Estatística descritiva do item 22, segundo modelo de Rasch. São Paulo (SP).

Item	Resposta	Escore	n	%	Média de <i>theta</i>	Média de erro-padrão	<i>Outfit</i>	Correlação item-total
foto 22	Tristeza	0	1	2	-0,12		0,1	-0,32
	Raiva	0	2	4	1,13	0,08	0,5	-0,19
	Surpresa	0	1	2	1,38		0,7	-0,09
	Alegria	0	1	2	1,38		0,7	-0,09
	Expectativa	0	6	11	2,31	0,16	1,8	0,13
	Neutro	0	2	4	2,69	0,18	2,5	0,15
	Aversão	1	16	29	1,92	0,24	1,3	-0,05
	Medo	1	26	47	2,09	0,18	1,0	0,11

Tabela 2. Propriedades psicométricas dos subtestes. São Paulo (SP).

Subtestes	Propriedade	Escore bruto	n	<i>Theta</i>	Erro do modelo	<i>Outfit</i>
Et_outr	Média	23,7	30	1,79	0,55	0,95
	Desvio-padrão	3,5		0,91	0,16	0,32
	Mínimo	12,0	30	-0,47	0,40	-0,9
	Máximo	29,0	30	3,81	1,04	1,90
Et_em_si	Média	32,0	42	1,80	0,49	0,93
	Desvio-padrão	6,8		1,35	0,19	0,45
	Mínimo	3,0	42	-3,34	0,35	0,28
	Máximo	41,0	42	4,41	1,03	3,13
EN_outr	Média	1,7	17	-2,64	0,92	0,91
	Desvio-padrão	1,1		0,62	0,17	0,55
	Mínimo	1,0	17	-3,08	0,57	0,41
	Máximo	5,0	17	-1,02	1,05	2,15
EN_em_si	Média	4,3	39	-3,30	0,82	0,68
	Desvio-padrão	5,8		1,56	0,31	0,51
	Mínimo	1,0	39	-4,91	0,37	0,03
	Máximo	36,0	39	3,24	1,24	2,13
t_din	Média	2,8	28	-2,68	0,74	0,93
	Desvio-padrão	1,6		0,73	0,20	0,54
	Mínimo	1,0	28	-3,68	0,47	0,27
	Máximo	7,0	28	-1,29	1,05	2,81

A precisão do subteste de percepção de emoções em si (Et_em_si) foi de 0,84 pelo método Rasch e de 0,88 pelo Coeficiente de Kuder-Richardson, considerados bons pelos critérios internacionais. A média de *theta* da amostra indica que este subteste também pode ser considerado com tendência a ser avaliado como fácil pelos sujeitos, e os valores de ajuste (*outfit*) encontraram-se dentro da faixa de adequação.

Em seguida, observa-se que a precisão do subteste de percepções neutras no outro (EN_outr) foi de 0,01 pelo método de Rasch e de 0,28 pelo Coeficiente de Kuder-Richardson, demonstrando que esse subteste é impreciso. Os valores médios de *theta* sugerem que a dimensão avaliada por meio dela é pouco frequente entre as pessoas da amostra. Os valores de ajuste (*outfit*) estão adequados.

A precisão do subteste de percepção de emoções neutras em si (EN_em_si) foi de 0,61 pelo método Rasch e de 0,93 pelo Coeficiente de Kuder-Richardson, considerado bastante aceitável. A média de *theta* da amostra também sugere que a dimensão avaliada por meio dela é pouco frequente entre as pessoas da amostra, implicando que os participantes tendem a relatar pouco que não sentem emoção ao observarem uma imagem que deveria despertar alguma emoção.

Os valores de ajuste (*outfit*) encontraram-se dentro dos parâmetros de adequação.

Finalmente, como o número de itens que deveriam compor o escore dinaceptivo era reduzido, optou-se por utilizá-lo para o subteste inteiro (t_din). A precisão desse subteste, segundo o modelo de Rasch, foi de 0,01, e segundo o Coeficiente de Kuder-Richardson foi de 0,46, indicando que a pontuação pelo escore dinaceptivo não é precisa. Os *thetas* dos participantes apresentaram-se baixos, o que significa que houve reduzida frequência de sujeitos assinalando emoções discrepantes. Os índices de ajuste (*outfit*) mostraram-se adequados.

O segundo bloco de resultados, referente à comparação entre os sistemas de pontuação por concordância com o consenso e com especialistas, é apresentado a seguir. A Tabela 3 mostra os coeficientes de correlação de Pearson entre as diversas pontuações obtidas.

Os padrões encontrados nos dados acima foram: alta correlação entre as pontuações por consenso e por especialistas (destacadas em negrito), baixa correlação entre a percepção de emoções em si e nos outros, e correlações negativas dessas pontuações com a percepção de emoções neutras e projetivas. A correlação entre percepção de emoções em si e nos outros indica

que esses dois subtestes estão relacionados, porém a correlação de magnitude moderada sugere que os subtestes mensuram aspectos diferentes do mesmo construto. Tal fato é coerente com a proposta de Mayer e Salovey (1997) de que a percepção de emoções envolve tanto a capacidade de identificar emoções em si mesmo como em outras pessoas ou ambientes. No MSCEIT (Mayer et al., 2002), as duas provas para avaliação da percepção de emoções contemplam apenas a identificação de emoções em outras pessoas ou em ambientes, mas negligenciam a percepção de emoções em si mesmo. Nesse sentido, o *Teste Informatizado de Percepção de Emoções em Fotos* (Miguel et al., 2006) pode ser considerado mais completo que o MSCEIT para avaliação de todas as facetas relacionadas à percepção de emoções, com boas propriedades psicométricas.

As correlações encontradas entre a percepção de emoções e percepções neutras no outro, assim como a correlação entre a percepção de emoções e percepções neutras em si, devem ser fruto do sistema de pontuação adotado. De acordo com o procedimento, só recebeu ponto em percepções neutras quem necessariamente não recebeu ponto nos subtestes de percepção de emoções.

Já as correlações verificadas entre o escore dinaceptivo e outras formas de pontuação merecem outras considerações. Embora o sistema de pontuação favoreça o aparecimento da correlação negativa, como ocorre entre as percepções de emoções e percepções neutras, nota-se uma grande diferença entre as correlações que o escore dinaceptivo estabeleceu com a percepção de emoções em si e nos outros. Enquanto a correlação com a percepção de emoções em si foi desprezível, a correlação com a percepção de emoções nos outros foi significativa e de alta magnitude. Esses dados sugerem que a falha na percepção de emoções

em si pode ser atribuída diretamente à incapacidade nessa habilidade. Contudo, o erro de percepção de emoções nos outros pode ser devido à interferência das dinacepções (Cattell, 1967), isto é, ao dinamismo emocional da própria pessoa que percebe, à medida que ela atribui aos outros as emoções que deveria perceber em si. Nesse sentido, o *Teste Informatizado de Percepção de Emoções em Fotos* (Miguel et al., 2006) acrescenta um subteste capaz de detectar a intensidade desses dinamismos.

Outro objetivo deste trabalho foi comparar a consistência das pontuações por concordância com especialistas e com o consenso, cujas correlações também se encontram na Tabela 3. A correlação entre os dois sistemas de pontuação no que se refere à percepção de emoções no outro foi de 0,82 ($p < 0,01$), e no que se refere à percepção de emoções em si foi de 0,86 ($p < 0,01$). Esses dados indicam uma convergência elevada e significativa dos dois métodos de pontuação, e são compatíveis com os resultados obtidos em estudos anteriores (Mayer, Salovey & Caruso, 2000; Mayer et al., 2003). Ao lado disso, esses resultados mostram que a estabilidade do construto analisado foi capaz de superar diferenças culturais e de instrumentos.

Apesar da convergência encontrada entre os dois sistemas de pontuação empregados, para verificar se havia alguma diferença entre os dois métodos de pontuação foi elaborado um gráfico de dispersão com as correlações item-total de cada sistema de pontuação para percepção de emoções em si e nos outros. O gráfico está apresentado na Figura 1, onde o eixo x representa os valores das correlações item-total de cada item segundo a pontuação por especialistas, e o eixo y apresenta os valores das correlações item-total segundo o consenso. Os itens relacionados à percepção de emoções em si estão representados por cz e os itens

Tabela 3. Coeficientes de correlação de Pearson entre subtestes. São Paulo (SP).

	Ct_outro	Ct_em_si	Et_outro	Et_em_si	EN_outro	EN_em_si	t_din
Ct_outro	1						
Ct_em_si	0,47**	1					
Et_outro	0,82**	0,47**	1				
Et_em_si	0,34*	0,86**	0,38**	1			
EN_outro	-0,50**	-0,38**	-0,41**	-0,26	1		
EN_em_si	-0,14	-0,60**	-0,14	-0,74**	0,16	1	
t_din	-0,67**	-0,30*	-0,71**	-0,24	0,13	0,11	1

* Correlação significativa no nível $p < 0,05$; ** Correlação significativa no nível $p < 0,01$.

relacionados à percepção de emoções nos outros estão representados por “f”.

A Figura 1 mostra que a atribuição de pontos por concordância com especialistas resulta em correlações item-total mais elevadas que aquelas da pontuação por concordância com o consenso, tanto para a percepção de emoção em si quanto nos outros. A média das correlações item-total para os 32 itens de percepção de emoções nos outros foi de 0,23 (desvio padrão - dp=0,13) para o consenso e 0,29 (dp=0,12) para especialistas. Em relação aos 45 itens relacionados à percepção de emoções nos outros, a média das correlações item-total foi de 0,25 (dp=0,18) para o consenso e 0,41 (dp=0,12) para especialistas. Para verificar

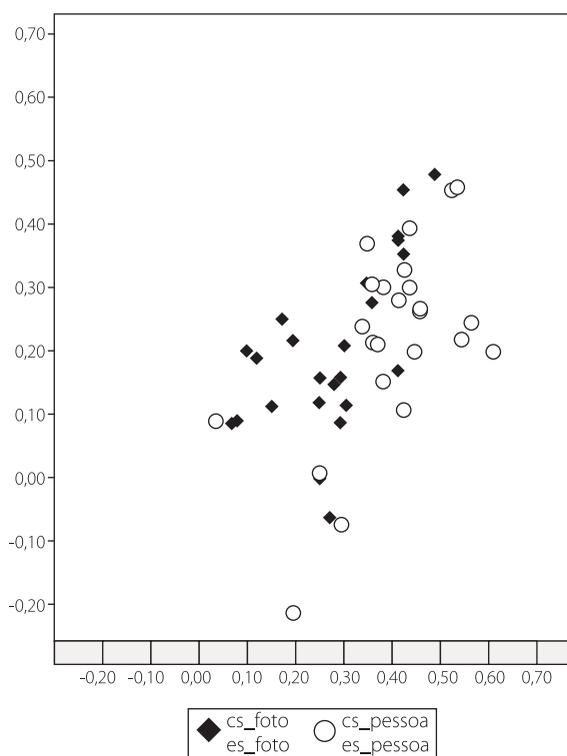


Figura 1. Correlações item-total entre as pontuações por consenso e por especialistas, nas fotos e nas pessoas. São Paulo (SP).

se as diferenças entre essas médias eram significativas, foi empregado o teste *t* de Student, cujos resultados estão apresentados na Tabela 4.

Os resultados do teste *t* de Student mostram que, embora os dois sistemas de pontuação estejam altamente correlacionados, houve diferença estatisticamente significativa entre as correlações item-total em favor das pontuações obtidas pela concordância com especialistas. Esses dados mostram que a opção pela atribuição de pontos pelo método de concordância com especialistas resulta em melhores índices de fidedignidade do instrumento.

De fato, o método de pontuação por concordância com o consenso não permite afirmar que um sujeito tenha maior ou menor acuidade perceptual das emoções, mas que ele concorda em maior ou em menor grau com a maioria das pessoas. Assim, supondo-se que, em um teste de inteligência, a dificuldade vá aumentando de item para item, é esperado que apenas uma minoria acerte os itens mais difíceis, que, segundo o critério de concordância com o consenso, receberiam pontuações menores do que a maioria que erra tais itens. Essa situação é inadmissível para um teste de aptidão (Bueno & Primi, 2003; Davies, Stankov & Roberts, 1998).

Considerações Finais

A proposta deste trabalho foi definir respostas corretas para subtestes de percepções em si e nos outros, com base na opinião de especialistas, e comparar suas propriedades psicométricas com as da atribuição de pontos por concordância com o consenso. A utilização do modelo de Rasch para identificação de itens inadequados foi fundamental para a definição das respostas corretas por especialistas. O conteúdo de itens que provocavam respostas inadequadas por sujeitos com alta habilidade ou respostas corretas por sujeitos com baixa habilidade pôde ser discutido quanto à sua adequação para a medição do construto e, quando

Tabela 4. Teste *t* de Student entre as correlações item-total das pontuações por concordância com o consenso e com especialistas. São Paulo (SP).

	Média	Desvio-padrão	t	gl	p	Correlações	p
cs_foto - es_foto	-0,05	0,11	-2,87	31	0,007	0,63	0,000
cs_pessoa - es_pessoa	-0,098	0,13	-4,95	44	0,000	0,53	0,000

necessário, quanto à permanência no teste final. Assim, alguns itens puderam ser melhorados, atribuindo-se mais de uma resposta correta, enquanto outros foram eliminados.

Essa metodologia resultou em um incremento significativo das propriedades psicométricas do teste relacionadas à fidedignidade. Isto pode ser observado pelo aumento médio das correlações item-total dos itens quando pontuados pelo sistema de concordância com especialistas, em detrimento da pontuação por concordância com o consenso. Ao lado disso, pelo menos um novo tipo de pontuação mostrou-se potencialmente interessante para avaliação de processos de percepção emocional: a pontuação em percepções dinceptivas (t_{din}). Esta faceta parece estar relacionada a um tipo de processamento em que as emoções funcionam de forma disruptiva, mas esta é uma questão que deve ser mais bem investigada em estudos posteriores.

O estudo também mostrou que o *Teste Informatizado de Percepção de Emoções em Fotos* é mais abrangente que as duas provas para avaliação da percepção de emoções que integram o MSCEIT, podendo ser utilizado para avaliação da capacidade de perceber emoções em si, em outras pessoas e em estímulos ambientais com confiabilidade. Portanto, este estudo não apenas constata as boas propriedades psicométricas do Teste Informatizado de Percepção de Emoções, como fornece indicadores significativos para o desenvolvimento de outras pesquisas, especialmente no sentido de obter um instrumento válido e fidedigno para avaliação dessa e de outras habilidades relacionadas à inteligência emocional.

Referências

- Bueno, J. M. H., & Primi, R. (2003). Inteligência emocional: um estudo de validade sobre a capacidade de perceber emoções. *Psicologia: Reflexão e Crítica*, 16 (2), 279-291.
- Cattell, R. B. (1967). Princípios de esquemas nos testes "projetivos" ou de má percepção da personalidade. In H. H. Anderson, G. L. Anderson & E. Bennett. *Técnicas projetivas do diagnóstico psicológico* (pp.69-111). São Paulo: Mestre Jou.
- Davies, M., Stankov, L., & Roberts, R. D. (1998). Emotional intelligence: in search of an elusive construct. *Journal of Personality and Social Psychology*, 75 (4), 989-1015.
- Ekman, P. (1965). Differential communication of affect by head and body cues. *Journal of Personality and Social Psychology*, 2 (5), 726-735.
- Lane, R. D. (2000). Levels of emotional awareness: neurological, psychological and social perspectives. In R. Bar-On & J. D. A. Parker (Orgs.), *The handbook of emotional intelligence: theory development, assessment, and applications at home, school, and in the workplace* (pp.171-191). San Francisco: Jossey-Bass.
- Lane, R. D., Fink, G. R., Chua, P. M., & Dolan, R. J. (1997). Neural activation during selective attention to subjective emotional responses. *Neuroreport*, 8 (18), 3969-3972.
- MacCann, C., Roberts, R. D., Matthews, G., & Zeidner, M. (2004). Consensus scoring and empirical option weighting of performance-based emotional intelligence (EI) tests. *Personality and Individual Differences*, 36 (3), 645-662.
- Mayer, J. D.; DiPaolo, M. T., & Salovey, P. (1990). Perceiving affective content in ambiguous visual stimuli: a component of emotional intelligence. *Journal of Personality Assessment*, 54 (3-4), 772-781.
- Mayer, J. D., & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence*, 22 (2), 89-113.
- Mayer, J. D., & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence*, 17 (4), 433-442.
- Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: implications for educators* (pp.3-31). New York: Basic Books.
- Mayer, J. D., Salovey, P., & Caruso, D. (1997). *Emotional IQ test* [CD ROM]. Needham, MA: Virtual Knowledge.
- Mayer, J. D., Salovey, P., & Caruso, D. (2000). Emotional intelligence meets standards for an intelligence. *Intelligence*, 27 (4), 267-298.
- Mayer, J. D., Salovey, P., & Caruso, D. R. (2002). *Mayer-Salovey-Caruso emotional intelligence test (MSCEIT): user's manual*. Toronto: MHS Publishers.
- Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion*, 3 (1), 97-105.
- Miguel, F. K., Primi, R., Muniz, M., Couto, G., & Noronha, A. P. (2006). *Teste informatizado de percepção de emoções em fotos*. Trabalho não-publicado, Universidade São Francisco, Itatiba.
- Plutchik, R. (2002). *Emotions and life: perspectives from psychology, biology and evolution*. Washington, DC: American Psychological Association.
- Primi, R., Bueno, J. M. H., & Muniz, M. (2006). Inteligência emocional: validade convergente e discriminante do MSCEIT com a BPR-5 e o 16PF. *Psicologia: Ciência e Profissão*, 26 (1), 26-45.
- Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, 9 (3), 185-211.
- Wright, B. D., & Stone, M. H. (2004). *Making measure*. Chicago: The Phanon Press.

Recebido em: 6/8/2007

Versão final reapresentada em: 28/5/2008

Aprovado em: 5/8/2008