

O FUNCIONAMENTO DIFERENCIAL DO ITEM NO TESTE PICTÓRICO DE MEMÓRIA

Fabián Javier Marín Rueda¹ – Universidade São Francisco

RESUMO

O objetivo do estudo foi verificar o ajuste do Teste Pictórico de Memória ao modelo Rasch, assim como também a presença de DIF em função do sexo das pessoas. Participaram da pesquisa 594 estudantes universitários de ambos os sexos e com idades variando de 18 a 56 anos. Foi aplicado de forma coletiva o Teste Pictórico de Memória. Quando analisado o instrumento com vistas a verificar seu ajuste ao modelo Rasch observou-se que os itens foram respondidos de acordo com o padrão esperado, apresentando, dessa forma, um bom ajuste. Quanto ao *oufit*, também sugeriu uma boa adequação ao modelo. Em relação ao DIF, os resultados mostraram que dos 55 itens estudados apenas cinco indicaram diferenciar o sexo, sendo que três favoreceram os homens e dois as mulheres. Com base nisso, concluiu-se que houve um certo equilíbrio nos vieses ocorridos para os homens e mulheres.

Palavras-chave: memória; funcionamento diferencial do item; validade.

THE DIFFERENTIAL ITEM FUNCTIONING IN THE TESTE PICTÓRICO DE MEMÓRIA

ABSTRACT

The objective of this study was to verify the adjustment of the Teste Pictórico de Memória to the Rasch model, as well to detect the presence of DIF in function of gender. Participated of the research 594 college students of both sex with ages between 18 and 56 years old. The Teste Pictórico de Memória was applied in an collective form. When analyzed the instrument had the objective to verify its adjustment to the Rasch model, was observed that items had been answered in accordance with the waited standard, presenting a good adjustment. About the *oufit*, it also suggested a good adequacy to the Rasch model. The aspects related to the DIF, had show that between 55 items studied only five had indicated differences of gender, whereas three had favored the men and two had favored women. Based on this information, the study concluded that was a certain balance in the bias occurred for the men and women. Keywords: memory, differential item functioning, validity.

Keywords: memory; differential item functioning; validity.

INTRODUÇÃO

Filósofos empíricos tais como John Locke, John Stuart Mill e Thomas Brown, ao longo dos séculos XVII e XIX, especularam sobre vários fatores que poderiam afetar o grau ou a força de associações particulares no âmbito do sistema psicológico. Segundo eles, a força das associações variaria de acordo com a vivacidade ou nitidez da experiência original, sua duração, sua frequência, e o interesse da pessoa para essa atividade. Foi proposto que o despertar de associações da memória (recordação) poderia ser de menor ou maior intensidade em razão da semelhança da pista estimulante para a memória, o quão recente fosse a experiência, a coexistência de poucos “associados alternativos” para a pista (chamados “interferência”), e as “diversidades temporárias de estado” (intoxicação, delírio, depressão). Tais conjecturas geraram muita pesquisa, e cada teoria

de memória lidou com esses fatores de formas diferentes (Bower & Hilgard, 1981).

Desde que a memória começou a ser estudada empiricamente, vários foram os pesquisadores e correntes teóricas que propuseram definir o que se entendia por esse construto. A primeira grande corrente foi a empírica, introduzindo a teoria de associação por contigüidade (Warren, 1921). Segundo essa corrente, idéias complexas seriam formadas na mente, conectando na memória idéias simples baseadas em sensações que seriam vivenciadas simultaneamente em tempo e/ou espaço. Em decorrência, a memória de que a o evento A foi vivenciado junto ou antecedendo imediatamente o evento B é registrada no banco de memória como uma associação da idéia A com a idéia B.

Segundo Warren (1921), para essa corrente, esse despertar das seqüências associativas da memória seria o método pelo qual as experiências passadas das pessoas causariam seus pensamentos posteriores e, dessa forma, o indivíduo poderia progredir de uma idéia para a outra. Essa noção básica foi elaborada para explicar a maneira pela

¹ *Contato:*

Rua José Marciano Filho, 9. Vila Cristo Redentor, Itatiba-SP.
CEP: 13251-420. E-mail: marinfabian@yahoo.com.br

qual os seres humanos desenvolvem expectativas coordenadas sobre propriedades de objetos, expectativas sobre seqüências causais de eventos, predições sobre eventos futuros, explicações de como ou por quê alguma coisa surgiu e planos de ação sobre resultados particulares.

Mas quando se fala em memória humana, deve-se dizer que o primeiro investigador experimental foi Ebbinghaus que, em 1885, se interessou em saber qual a quantidade de informação que as pessoas poderiam se lembrar de um teste de memória, imediatamente após a apresentação da informação (Tulving & Craik, 2000). Segundo Tulving e Craik (2000), foi Ebbinghaus quem inventou a noção da sílaba sem sentido (como DAX, QEH) para fornecer a si próprio materiais de aprendizagem de dificuldade homogênea, evitando, dessa forma, a variabilidade de palavras familiares. Ele se “auto-ensinou” lendo em voz alta listas seriais de 6 a 20 sílabas com um metrônomo, tentando posteriormente recitar as séries de memória.

Ainda de acordo com os autores, Ebbinghaus introduziu muitas idéias e métodos importantes. Mensurou a dificuldade para aprender uma lista (números ou palavras) a partir do número de vezes de estudo exigidas para conseguir sua recitação sem erro, notando que a dificuldade aumentava de forma desproporcional e em função da extensão da lista a ser aprendida. Dessa forma, introduziu a idéia de “graus de aprendizagem” mensuráveis, verificando os registros na re-aprendizagem de uma lista que ele tinha aprendido anteriormente. O paradigma de Ebbinghaus e o fenômeno que ele descobriu, suas idéias e seus métodos, tiveram influência ao longo do século XX nas pesquisas sobre memória humana.

Investigações subseqüentes propuseram outros paradigmas e testaram muitas variáveis que determinaram a performance da memória em diferentes contextos. As memórias começaram a ser testadas tanto por lembrança, reconhecimento, reconstrução, assim como por uma variedade de medidas indiretas. A natureza dos materiais poderia ser variada, como também poderia ser variada a forma de apresentação, estratégias que os sujeitos usariam para estudá-las, expectativas a respeito do teste de memória e relações entre vários conjuntos de materiais a serem aprendidos. Como resultado, uma enorme quantidade de informação empírica foi acumulada sobre como os humanos aprendem em dadas situações, além de muitas hipóteses teóricas terem sido aventadas e testadas para integrar as

diferentes definições e propostas para entender a memória humana (Tulving & Craik, 2000).

Nas décadas de 50 e 60 do século XX houve um grande desenvolvimento da pesquisa sobre memória na corrente cognitivista. Os psicólogos cognitivistas foram inspirados pela teoria de estímulos-resposta (S-R), que concebia a memória como o processo de codificar e decodificar uma informação, atribuindo à memória o conceito de “capacidade limitada” (Garner, 1962). Essa corrente forneceu argumentos para uma visão do “processamento de informação” dos seres humanos. As pessoas eram vistas como levando informação para um sistema perceptivo, atendendo seletivamente partes dessa informação, codificando ou transformando-a por meio de suas habilidades cognitivas, armazenando-a na memória, e devolvendo-a quando uma pista de devolução apropriada fosse ativada. A análise da percepção, atenção, memórias sensoriais imediatas, memória de curto prazo e a estrutura da memória de longo prazo se tornou um tópico proeminente.

Broadbent (1958) propôs o primeiro modelo estrutural do processamento da informação no sistema cognitivo humano. Ele representa o primeiro diagrama que mostraria de que forma a informação fluiria no sistema de processamento da informação e o que aconteceria com a informação percebida e não percebida. Outros psicólogos da época, interessados mais diretamente no estudo da memória, propuseram também modelos estruturais semelhantes ao de Broadbent para tentar dar sentido aos resultados de suas investigações. Dentre esses modelos, o que mais influenciou nas pesquisas posteriores sobre a memória humana foi o proposto por Atkinson e Shiffrin (1968). Tal modelo é denominado estrutural ou modal, pois dá ênfase na existência de três estruturas ou diferentes depósitos da memória, quais sejam, os registros sensoriais, a memória de curto prazo e a memória de longo prazo.

O registro sensorial seria um depósito que reteria temporariamente a informação, antes de transferi-la para a memória de curto prazo, na qual essa informação seria analisada e codificada (fase de aquisição). Essa aquisição dependeria da atenção, da percepção e da codificação do material a ser apreendido. Se essa informação estocada na memória de curto prazo não for utilizada pelo indivíduo após a fase de aquisição, pode ser perdida após um período que variaria de 15 segundos a 1 minuto. Também, algumas dessas informações presentes na memória de curto prazo poderiam ser

transferidas para a memória de longo prazo, a qual é considerada um depósito permanente da informação (Atkinson & Shiffrin, 1968).

Com o advento da abordagem cognitiva, diferentes questões sobre memória foram feitas, levando a vários paradigmas, os quais se tornaram populares porque forneciam um meio para estudar diferentes tipos de processos ou de memória. Esses diferentes tipos de memória foram lembrança livre, memória de reconhecimento, semântica, assim como testes de memória diretos e indiretos (Tulving & Craik, 2000).

Dessa forma, pode-se dizer que mesmo com inúmeras variações e elaborações, o experimento de memória contemporâneo consiste de três fases, quais sejam, uma fase de estudo ou codificação, na qual o material é apresentado ao sujeito, um intervalo de retenção, e finalmente uma fase de devolução ou teste, na qual o sujeito tenta responder a uma questão que envolve o uso da informação inicialmente estudada. Os diferentes métodos de memória, de Ebbinghaus até os dias de hoje, podem ser caracterizados em termos das condições que eles estabelecem para cada uma dessas três fases. A estratégia de pesquisa fundamental tem sido variar as condições em cada uma das fases, e os paradigmas de pesquisa de memória têm consistido em um número crescente de variações nas condições de cada fase.

É importante salientar que a mudança na forma de abordagem não trouxe modificações radicais no método, mas deu-lhe um novo propósito, produzindo outra ênfase e o surgimento de muitos procedimentos novos. Cada uma das três fases passou a ser vista como um conjunto de operações complexas a ser entendido em termos de um processador de informação ativo. Além disso, o foco desses métodos mudou o entendimento da interação com outras variáveis em cada uma das fases. Por exemplo, ao invés de perguntar como diferentes processos de codificação influenciariam os níveis de performance (causa e efeito), a questão de maior interesse se tornou como a lembrança poderia ser determinada pela interação de processos de codificação particulares com diferentes condições no momento da devolução (processo de resposta).

Com base no até aqui descrito, pôde-se pensar na possibilidade de verificar se o processo de resposta (devolução) num teste de memória, mais especificamente o Teste Pictórico de Memória (Rueda & Sisto, no prelo) sofreria a influência da variável sexo. Para tal, valeu-se de um método

estatístico muito utilizado atualmente, o funcionamento diferencial do item (DIF). Vale ressaltar que esse tipo de estudo pode ser considerado um tipo de evidência de validade, uma vez que permitiria verificar se os itens de um teste sofreriam influência de algum tipo de variável externa, no caso específico deste estudo, o sexo dos indivíduos. Dentro desse contexto, deve-se considerar que um bom modelo de medida deve requerer, pelo menos, que um teste válido satisfaça três aspectos. Um deles é que uma pessoa mais capaz (ou com maior quantidade de um traço) tenha maior probabilidade de acertar um item do que uma pessoa menos capaz. Outro aspecto seria que toda pessoa tenha mais possibilidade de acerto em um item considerado fácil do que em um difícil. E, por fim, que essas condições não devem estar relacionados à raça, sexo, entre outras características que possam interferir em seu desempenho (Sisto, 2006a).

Levando em conta essas condições, a compreensão de que a dificuldade de um item é uma propriedade intrínseca e específica do item em qualquer circunstância, não poderia haver qualquer interferência de sua cultura, sexo e outras características, sendo congruente a qualquer população em particular. Analogamente, a habilidade de uma pessoa é definida como uma característica dela sem referência a qualquer conjunto particular de itens, tendo como concepção uma interação entre esses dois elementos, trazendo como consequência dois parâmetros: a habilidade da pessoa e a dificuldade do item (Sisto, 2006a; 2006b).

Para Andriola (2001), quando outros aspectos produzem aumento ou diminuição da dificuldade do item, que não seja a interação entre ele e a pessoa, é possível que haja viés na medida, ultimamente estudado sob a denominação de funcionamento diferencial do item (DIF). Segundo Sisto (2006a; 2006b), os procedimentos para detectar o DIF têm em comum o fato de utilizar os resultados globais do teste e se divide em duas categorias, dependendo se o critério de viés é externo ou interno. Em qualquer uma das situações o objetivo é distinguir os itens que definem o traço ou se o item está enviesado em alguma forma por algum subgrupo.

Uma das propostas para detectar esses possíveis vieses nos itens de um teste foi realizada por Georg Rasch, que em 1960 forneceu um repensar sobre o problema da medida ao abordar adequadamente a maioria dos problemas em relação

às deficiências da análise tradicional do item. Seu modelo estocástico descreve a probabilidade de acerto de uma pessoa em um item em função somente da sua habilidade e da dificuldade do item. Com base na exigência tradicional de que uma medida deve ser constituída por um conjunto de itens homogêneos, unidimensionalmente relacionada ao traço a ser medido, Rasch derivou seu modelo de medida na forma de uma expressão logística e demonstrou que, nessa forma, os parâmetros do item e da pessoa são estatísticas separáveis.

Segundo Sisto (2006a; 2006b) o modelo logístico de Rasch pôde incorporar a maior parte do trabalho prévio sobre viés, porque começou com suposições similares de medida. Assumindo a definição tradicional de viés do item, no sentido de que um item é enviesado se medir diferentemente distintos grupos, esse modelo avalia tal efeito por meio da análise dos resíduos. Mas como seus procedimentos são extensões racionais do modelo, a análise do viés pode ser averiguada de maneira mais sistemática e mais integrada do que tem sido feito.

Talvez a melhor justificativa para a necessidade de se avaliar o funcionamento diferencial dos itens em um teste de memória, seja que, por um lado, a literatura reconhece que a memória é essencialmente importante em vários aspectos da vida dos indivíduos e, por outro lado, têm-se questionado se a memória seria influenciada por variáveis externas, como por exemplo, o sexo, o nível sócio-econômico, ou pelo fato da pessoa ter sido mais ou menos estimulada academicamente ao longo da sua vida (Tulving & Craik, 2000).

O Teste Pictórico de Memória (Rueda & Sisto, 2007) foi construído como uma alternativa para avaliar a capacidade do indivíduo recuperar uma informação num curto período de tempo. O Manual oferece alguns indicadores de evidência de validade. Num primeiro momento foi verificada evidência de validade em relação ao processo de respostas. Nesse sentido, os itens dos três agrupamentos que compõem o teste (céu, água e terra) foram agrupados. Com essas categorias calculou-se suas médias ponderadas, em razão das diferentes quantidades de itens em cada ambiente. Feito isso, realizou-se uma Análise de Variância (ANOVA), sendo que a prova de *Tukey* diferenciou os três grupos [$F(2, 642) = 256,843, p=0,000$]. Com base nessa distinção concluiu-se que há diferentes processos de resposta no teste, o que pode ser considerado uma evidência de validade.

Também foi verificada evidência de validade desenvolvimental, observando que com o aumento da idade, diminui a capacidade do indivíduo recuperar uma informação num curto período de tempo. Por fim, evidência de validade com o Teste de Raciocínio Inferencial Sisto (2006c) também foi estudada. Os resultados mostraram correlações positivas e significativas, com coeficientes moderados. Quando controlado o efeito da idade os coeficientes de correlação se mantiveram, assim como também os níveis de significância e as tendências das correlações. Dessa forma, detectou-se que há uma associação entre a memória avaliada pelo Teste Pictórico de Memória e a inteligência medida pelo Teste de Raciocínio Inferencial.

Assim, dentro desse contexto, este estudo tem como objetivo verificar o ajuste do Teste Pictórico de Memória ao modelo Rasch, assim como também se os itens sofrem um viés relacionado ao sexo, pois a possibilidade de subavaliar sistematicamente as capacidades em razão do gênero, justifica a importância de estudo para neutralizar esse tipo de efeito. Essa análise será feita por meio do modelo logístico proposto por Rasch, que possibilita identificar itens enviesados, considerando também a dificuldade do item e a habilidade da pessoa.

MÉTODOS

Participantes

Participaram da pesquisa 594 estudantes universitários dos estados de São Paulo e Minas Gerais, com idades variando de 18 a 56 anos ($M=21,03, DP=6,52$), sendo 387 mulheres (65,15%) e 207 homens (38,85%).

Instrumento

O Teste é composto por uma figura com vários desenhos e detalhes que podem ser agrupados em três categorias, quais sejam, itens que pertencem e podem ser encontrados na categoria Água (pato, pescador, jet-ski, dentre outros); itens referentes à categoria Céu (helicóptero, balão, nuvem, sol, por exemplo) e itens que podem ser localizados na categoria Terra (casa, cadeira, mesa, fogueira, dentre outros). Para responder ao teste a pessoa deve visualizar a figura durante um minuto e, em seguida, deve lembrar a maior quantidade de desenhos e detalhes possíveis e escrevê-los na folha de resposta do teste. A pontuação pode variar de 0 a 55, sendo que é atribuído 1 ponto para cada item lembrado pelo indivíduo.

Procedimento

Após assinatura do Termo de Consentimento Livre e Esclarecido por parte dos respondentes, o instrumento foi aplicado de forma coletiva. Para tal, foi projetado o desenho na tela por durante 1 minuto, e após esse período de tempo foi dado aos indivíduos 2 minutos para escreverem a maior quantidade de desenhos e detalhes do desenho. A aplicação não excedeu a 30 pessoas por grupo. O tempo de aplicação foi de aproximadamente 5 minutos.

RESULTADOS

O programa estatístico usado para fazer a análise pelo modelo Rasch foi o Winsteps, Nesse

programa, tecnicamente, o modelo estima quanto de DIF é acrescentado (positiva e negativamente) ao item e, por meio da prova *t* de Student, calcula se a diferença pode ser atribuída ao acaso ou não. Em geral, um valor de *t* maior do que dois é considerado significativo. No entanto, Draba (1977) concordando com Bonferroni, considera que o patamar de 2,4 é um ponto bom para a significância estatística, quando se está analisando mais de 20 itens.

O instrumento foi analisado com vistas a verificar seu ajuste ao modelo Rasch. Para isso foram considerados todos os itens indistintamente. Os resultados dessa análise constam da Tabela 1. São oferecidos os resultados em relação a dificuldades dos itens, erros padrão, *infits*, *oufits*.

Tabela 1. Índices estatísticos do ajuste dos itens ao modelo *Rasch*: dificuldades dos itens, erros padrão, *infits*, *oufits*.

<i>Item</i>	<i>Dificuldade</i>	<i>Erro</i>	<i>Infít</i>	<i>Outfít</i>
Linha da pipa	4,06	0,58	1,01	2,38
Pessoa em balão	2,58	0,28	1,02	1,43
Céu	1,94	0,21	1,03	1,36
Carro	-0,95	0,08	1,12	1,19
Balão	-3,07	0,11	1,08	1,19
Caminho	0,73	0,13	1,03	1,18
Porta	2,58	0,28	0,99	1,17
Árvore	-1,16	0,08	1,09	1,14
Parque	1,34	0,16	1,03	1,14
Gramma	5,17	1,00	1,00	1,13
Balde	3,21	0,38	1,01	1,12
Chaminé	-0,11	0,10	1,04	1,10
Pessoa em trampolim	0,50	0,12	1,03	1,10
Janela	2,18	0,23	0,99	1,09
Pára-quedas	-0,54	0,09	1,01	1,06
Pessoa em pára-quedas	1,78	0,19	1,01	1,06
Foguete	-1,91	0,09	1,03	1,03
Deck	0,39	0,11	0,96	1,03
Trampolim	-0,50	0,09	1,03	1,02
Pára-quedista	1,54	0,18	1,01	1,02
Escorregador	-1,20	0,08	1,02	1,02
Avião	-3,38	0,12	1,02	1,02
Casa	-4,21	0,16	1,00	1,02
Água	-0,73	0,09	1,01	1,01
Criança	-1,25	0,08	1,01	1,01
Lancha	-0,17	0,10	1,01	0,99
Rabiola	4,47	0,71	1,00	0,74
Barraca	-0,86	0,09	1,00	0,98
Pássaro	-1,21	0,08	1,00	1,00
Balanço	-1,89	0,09	1,00	1,00
Terra	3,21	0,38	0,99	0,65
Mesa	-0,24	0,10	0,99	0,98

Continuação da Tabela 1				
Nuvem	-2,72	0,10	0,99	0,96
Gangorra	-1,04	0,08	0,99	0,99
Jet-ski	-0,37	0,09	0,99	0,96
Surfista	-0,79	0,09	0,99	0,98
Vara de pescar	0,94	0,14	0,99	0,94
Prancha	1,29	0,16	0,98	0,82
Vela	1,19	0,15	0,98	0,82
Cadeira	-0,23	0,10	0,98	0,95
Barco	-2,03	0,09	0,98	0,97
Pescador	-1,87	0,08	0,98	0,97
Onda	1,00	0,14	0,98	0,89
Fogueira	-1,09	0,08	0,98	0,96
Sol	-2,53	0,09	0,97	0,94
Helicóptero	-2,44	0,09	0,96	0,97
Cesta de basquete	-0,51	0,09	0,95	0,97
Poço	-1,12	0,08	0,96	0,94
Pato	-0,56	0,09	0,96	0,93
Fumaça	0,22	0,11	0,95	0,87
Bola	-0,19	0,10	0,94	0,90
Peixe	0,10	0,10	0,94	0,85

Uma síntese dos resultados referentes aos itens pode ser observada na Tabela 2, que indicaram uma média do *infit* de 1,00 (DP=0,03). Assim, é possível interpretar que, os itens foram respondidos de acordo com o padrão esperado, que é de 1,00. Em uma análise mais detalhada, observa-se que os valores *infit* ficaram entre 0,94-1,12, enquadrando-se no intervalo 0,70-1,30, normalmente considerado como critério de bom ajuste. Dessa forma, interpretou-se que a variabilidade do *infit* dos itens esteve entre níveis esperados e que o ajuste está adequado.

Tabela 2. Parâmetros de ajuste dos itens

Parâmetros	Itens		
	<i>Infit</i>	<i>Outfit</i>	<i>erro</i>
Média	1,00	1,03	0,17
D.P.	0,03	0,23	0,17
Máximo	1,12	2,38	1,00
Mínimo	0,94	0,65	0,08
1,3 < > 1,5	0 (0,00%)	2 (1,10%)	
1,5 < > 2,0	0 (0,00%)	0 (0,00%)	
< 2,0	0 (0,00%)	1 (0,55%)	

Em relação ao *outfit* seu valor médio foi de 1,03 (DP=0,23), sugerindo uma boa adequação ao modelo. No entanto, seus valores tiveram uma variação entre 0,74-2,38. O item com valor

discrepante foi a “*linha da pipa*”. Esse resultado indicou que esse item foi lembrado por pessoas cujas habilidades não sugeriram a possibilidade de lembrá-lo. Além desse item, na “*pessoa em balão*” e no “*céu*” observou-se que seu *outfit* foi superior a 1,30, mas não superior a 1,5, valor esse que Linacre (2002) considera como limite máximo para se aceitar um item.

Em relação às pessoas os dados estatísticos estão na Tabela 3. No que concerne ao *infit*, observou-se que 9 pessoas (1,42%) forneceram escolhas desajustadas, pois maior que 1,50 (Linacre,2002). De certa forma, as respostas dessas pessoas não foram compatíveis com o esperado na escala em razão de suas habilidades para se lembrar. Por sua vez, os dados do *outfit* registraram que 32 (5,06%) pessoas com respostas inesperadas. De forma geral, os percentuais de desajuste podem ser considerados baixos.

Destaca-se que a maioria das pessoas apresentou o padrão esperado que seria de 1,00, pois a média do *infit* foi 1,00 (DP=0,19) e a do *outfit* de 1,01 (DP=0,89), com uma média de erro de 0,36 (DP=0,03), conforme pode ser observado na Tabela 3. Na Tabela 4 pode ser observado o agrupamento dos itens e das pessoas no teste, onde a letra M, ao lado da divisória das informações das pessoas e dos itens, indica a localização da média.

Tabela 3. Parâmetros de ajuste das pessoas

Parâmetros	Itens		
	<i>Infit</i>	<i>Outfit</i>	<i>erro</i>
Média	1,00	1,01	0,36
D.P.	0,19	0,89	0,03
Máximo	1,84	9,90	0,64
Mínimo	0,59	0,38	0,34
1,3 < > 1,5	18 (2,84%)	8 (1,26%)	
1,5 < > 2,0	9 (1,42%)	1 (0,16%)	
< 2,0	0 (0,00%)	31 (4,90%)	

Tabela 4. Mapa de itens e pessoas.



Por essa Tabela observou-se que a média das pessoas ficou muito abaixo da média dos itens. Esse dado facilitou a interpretação de que a quantidade de itens lembrados foi muito baixa.

Avaliadas essas propriedades gerais do instrumento, procedeu-se ao estudo do DIF. Os resultados encontram-se na Tabela 5.

Tabela 5. Acréscimos de DIF por sexo, mudança e valores de t por item.

Sexo	DIF acrescido	Sexo	DIF acrescido	Mudança	t	item
M	-0,01	F	0,00	-0,01	-0,06	água
M	-0,01	F	0,00	-0,02	-0,09	pato
M	-0,17	F	0,07	-0,23	-1,03	peixe
M	-0,27	F	0,12	-0,39	-1,62	deck
M	-0,30	F	0,11	-0,41	-2,08	pescador
M	0,36	F	-0,12	0,47	1,39	vara
M	0,32	F	-0,10	0,42	0,37	linha da vara
M	-0,05	F	0,02	-0,07	-0,38	barco
M	-0,41	F	0,20	-0,61	-1,93	vela
M	-0,32	F	0,13	-0,45	-2,23	jet-ski
M	-0,34	F	0,14	-0,49	-2,48	trampolim
M	0,35	F	-0,12	0,47	1,65	pessoa em trampolim
M	-0,34	F	0,15	-0,49	-2,35	lança
M	-0,09	F	0,03	-0,12	-0,40	onda
M	-0,03	F	0,01	-0,04	-0,21	surfista
M	0,32	F	-0,10	0,43	1,10	prancha
M	-0,44	F	0,23	-0,67	-1,56	céu
M	0,18	F	-0,07	0,25	1,35	arco-iris
M	0,22	F	-0,09	0,31	1,30	balão
M	-0,12	F	0,05	-0,16	-0,27	pessoa em balão
M	0,00	F	0,00	0,01	0,03	helicóptero
M	-0,07	F	0,02	-0,09	-0,33	avião
M	-0,20	F	0,08	-0,27	-1,40	para-quedas
M	-0,04	F	0,01	-0,05	-0,12	pessoa em paraquedas
M	-0,41	F	0,20	-0,61	-1,69	paraquedista
M	0,12	F	-0,05	0,17	0,82	sol
M	0,47	F	-0,20	0,67	3,19	nuvem
M	0,03	F	-0,01	0,04	0,23	pássaro
M	-0,21	F	0,07	-0,28	-1,45	foguete
M	-0,45	F	0,24	-0,69	-0,90	terra
M	0,13	F	-0,05	0,18	0,94	árvore
M	-0,06	F	0,02	-0,09	-0,46	fogueira
M	0,12	F	-0,05	0,16	0,84	barraca
M	-0,13	F	0,05	-0,18	-0,95	carro
M	0,33	F	-0,12	0,46	2,37	gangorra
M	0,27	F	-0,10	0,38	1,99	escorregador
M	0,21	F	-0,08	0,28	1,51	balanço
M	-0,72	F	0,45	-1,18	-3,60	parque
M	0,05	F	-0,02	0,07	0,34	bola
M	-0,15	F	0,06	-0,20	-1,03	cesta de basquete
M	0,47	F	-0,16	0,62	2,68	cadeira
M	0,48	F	-0,16	0,64	2,79	mesa
M	-0,01	F	0,00	-0,01	-0,07	poço

Continuação da Tabela 5						
M	-0,04	F	0,01	-0,05	-0,06	balde
M	0,14	F	-0,05	0,20	1,05	criança
M	-0,16	F	0,06	-0,22	-1,13	pipa
M	-0,20	F	0,08	-0,28	-0,23	linha da pipa
M	mínimo	F	-0,32			rabiola
M	-0,52	F	0,14	-0,66	-1,55	casa
M	-0,14	F	0,06	-0,20	-0,40	janela
M	-0,12	F	0,05	-0,16	-0,27	porta
M	0,23	F	-0,08	0,32	1,39	chaminé
M	0,20	F	-0,07	0,27	1,07	fumaça
M	-0,27	F	0,12	-0,39	-1,44	caminho
M	mínimo	F	-0,32			grama

Dos 55 itens estudados, apenas 5 deles indicaram diferenciar o sexo. No caso das mulheres, os itens que as favoreceram foram “trampolim” e “parque”. Já para os homens, eles foram favorecidos pelos itens “nuvem”, “cadeira” e “mesa”. Pelo fato de três itens terem sido mais fáceis para os homens e dois para as mulheres, concluiu-se que houve um certo equilíbrio nos vieses ocorridos para os homens e mulheres. Como uma certa quantidade de DIF sempre está presente, julgou-se que não seria necessário a substituição dos itens. Com base nisso, esta pesquisa forneceu evidência de validade para o Teste Pictórico de Memória (Rueda & Sisto, 2007), uma vez que os itens se apresentaram sem vieses.

À GUIA DE CONCLUSÃO

O presente estudo foi proposto com o objetivo de verificar o ajuste do Teste Pictórico de Memória (Rueda & Sisto, 2007) ao modelo Rasch e se os itens do apresentam algum tipo de viés em função do sexo da pessoa que responde. Este tipo de estudo é de suma importância na sociedade atual, uma vez que há aquelas pessoas que defendem direitos iguais para ambos os sexos. Assim, a criação de instrumentos que controlem esse efeito é de suma importância, uma vez que o que deve ser avaliado é a capacidade do indivíduo *per se* e não em função de ser homem ou mulher.

Nesse sentido, o Teste Pictórico de Memória mostrou-se uma medida da capacidade do indivíduo devolver uma informação num curto período de tempo isenta desse problema, o que evidencia uma validade de construto relacionada ao item. A pontuação obtida pode ser interpretada sem o risco de se estar valorizando característica de um sexo em detrimento do outro.

Além disso, os itens foram respondidos de acordo com o padrão esperado quando analisado o teste pelo modelo Rasch, e o número de pessoas que apresentou respostas desajustadas de acordo com o modelo foi muito pequeno. Ressalta-se a importância de realização de estudos desta natureza, com o intuito de minimizar erros na avaliação realizada com os indivíduos.

REFERÊNCIAS

- Andriola, W. B. (2001). Descrição dos principais métodos para detectar o funcionamento diferencial dos itens (DIF). *Psicologia: Reflexão e Crítica*, 14 (3), 643-652.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory. A proposed system and its control processes. Em K. W. Spence & J. T. Spence (Orgs.). *The psychology of learning and motivation*, (pp. 89-195). New York: Academic Press.
- Bower, G. H., & Hilgard, E. (1981). *Theories of Learning*. New York: Englewood Cliffs.
- Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon Press.
- Draba, R. E. (1977). *The Identification and Interpretation of Item Bias*. *Rasch Measurement Transactions*, MESA Memorandum no. 25, Disponível em: <http://www.rasch.org/rmt/rmt122m.htm>. Consultado em 17/04/2004.
- Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: John Wiley Sons.
- Linacre J.M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16(2), 887.

- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danmarks Paedagogiske Institut.
- Rueda, F. J. M., & Sisto, F. F. (2007). *Teste Pictórico de Memória. Manual*. São Paulo: Vetor Editora Psicopedagógica Ltda.
- Sisto, F. F. (2006a). Estudo do funcionamento diferencial de itens para avaliar o reconhecimento de palavras. *Avaliação Psicológica*, 5(1), 1-10.
- Sisto, F. F. (2006b). O funcionamento diferencial do item. *Psico-USF*, 11(1), 35-43.
- Sisto, F. F. (2006c). *Teste de Raciocínio Inferencial*. São Paulo: Vetor Editora Psicopedagógica Ltda.
- Tulving, E., & Craik, F. I. M. (2000). *The Oxford Handbook of Memory*. Oxford: University Press.
- Warren, H. C. (1921). *A history of the association philosophy*. New York: Charles Scribner's Sons.

Recebido em Maio de 2007

Reformulado em Agosto de 2007

Aceito em Outubro de 2007

SOBRE O AUTOR:

Fabián Javier Marín Rueda: psicólogo, perito e mestre em avaliação psicológica, e doutorando do programa de pós graduação *Stricto Sensu* em psicologia da Universidade São Francisco. Bolsista CAPES.