



Psicologia: Reflexão e Crítica

ISSN: 0102-7972

prcrev@ufrgs.br

Universidade Federal do Rio Grande do

Sul

Brasil

Cecilio Fernandes, Débora; Prieto, Gerardo; Delgado, Ana R.  
Construção e análise pelo modelo de Rasch de dois testes computadorizados de memória de  
reconhecimento

Psicologia: Reflexão e Crítica, vol. 28, núm. 1, enero-marzo, 2015, pp. 49-60

Universidade Federal do Rio Grande do Sul

Porto Alegre, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=18838355006>

- ▶ Como citar este artigo
- ▶ Número completo
- ▶ Mais artigos
- ▶ Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica

Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal  
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

# Construção e Análise pelo Modelo de Rasch de Dois Testes Computadorizados de Memória de Reconhecimento

## *Construction and Analysis by the Rasch Model of Two Computerized Recognition Memory Tests*

Débora Cecilio Fernandes<sup>\*, a</sup>, Gerardo Prieto<sup>b</sup> & Ana R. Delgado<sup>b</sup>

<sup>a</sup>Universidade do Vale do Sapucaí, Pouso Alegre, MG, Brasil & <sup>b</sup>Universidad de Salamanca, Salamanca, España

### **Resumo**

Os objetivos deste estudo foram desenvolver uma versão computadorizada para um teste de reconhecimento associativo, desenvolver itens para a construção e informatização de um teste de reconhecimento de itens, utilizando o modelo de Rasch. Também foi objetivo deste estudo testar as hipóteses dos níveis de dificuldade dos itens e buscar evidência de validade para os instrumentos, averiguando se as mulheres obtêm desempenho superior ao dos homens. Participaram 217 estudantes universitários, homens e mulheres, com idade média de 25 anos. Os resultados indicaram que os testes se ajustaram ao modelo de Rasch, apresentaram índices de fidedignidade adequados e não apresentaram funcionamento diferencial dos itens (DIF). Obteve-se empiricamente a ordem esperada de dificuldade das categorias de itens do teste individual e uma alta correlação entre a ordem de dificuldade esperada dos pares e a ordem obtida dos pares. As mulheres apresentaram um desempenho significativamente superior ao dos homens, que não se deve ao DIF.

*Palavras-chave:* Memória de reconhecimento, lembrança, familiaridade, modelo de Rasch, avaliação.

### **Abstract**

The objectives of this study were to develop items for a computerized version of an associative recognition memory test and to develop items for the construction and computerization of an individual recognition memory test, using the Rasch model. This study also aimed to test the hypothesis related to the item difficulty levels and to search for validity evidences for the instruments investigating if there are differences between male and female performance. Two hundred and seventeen college students, male and female, with average age of 25 years, participated in the study. Results indicated that the tests fit the Rasch model, presented adequate reliability index and did not present differential item functioning (DIF). The expected difficulty order of item categories from the individual subtest as well as a large correlation between the pairs expected difficulty order and the pairs obtained order were found. Women presented significantly higher performance than men and it is not due to DIF.

*Keywords:* Recognition memory, recollection, familiarity, Rasch model, assessment.

A memória de reconhecimento é compreendida como a nossa capacidade para decidir que já estivemos em contato com um estímulo anteriormente (Yonelinas, 1994). Muitas teorias da memória de reconhecimento consideravam e ainda consideram que o reconhecimento era medido, exclusivamente, pela familiaridade (Mandler, 2008), ideia representada pela Teoria da Detecção do Sinal (TDS), que entende o processo de familiaridade como a intensidade ou a força de uma lembrança que é usada como base para o reconhecimento (Gillund & Shiffrin, 1984; Hintzman, 1986; Murdock, 1997; Yonelinas, 1994).

Os modelos de processo duplo, cujo principal representante é o modelo híbrido denominado de processo duplo

de detecção do sinal (PDDS), proposto por Yonelinas e colaboradores (Yonelinas, 1994, 1997; Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996) assumem que a memória de reconhecimento depende da contribuição relativa de dois processos subjacentes distintos: a familiaridade e a lembrança. A familiaridade é entendida como um processo rápido e automático, com pouca ou nenhuma demanda de atenção e é sensível às mudanças perceptivas dos estímulos estudados. A lembrança é um processo semelhante à lembrança livre, ou seja, um processo de busca das informações, caracterizado como mais lento, intencionado, controlado e com demandas de atenção (Defeyter, Russo, & McPartlin, 2009).

Estudos indicam que ambos os processos são disociáveis e podem atuar em conjunto para facilitar as respostas de reconhecimento, no entanto, a participação e a contribuição de cada processo dependem do tipo de tarefa de reconhecimento, conforme revelaram as pesquisas de

\* Endereço para correspondência: Av. Prefeito Tuany Toledo, 470, Fátima I, Pouso Alegre, MG, Brasil 37550-000. E-mail: debora.cecilio@gmail.com, gprieto@usal.es e adelgado@usal.es

Aggleton e Brown (2006), Parks e Yonelinas (2007) e Woodruff, Hayama e Rugg (2006). De forma conjunta, esses estudos mostram que a familiaridade contribui, principalmente, com as tarefas de reconhecimento de itens, enquanto que a lembrança colabora, em maior medida, com as tarefas de reconhecimento associativo (Wheeler & Buckner, 2004; Woodruff et al., 2006; Yonelinas, 2002). O reconhecimento de itens (ou reconhecimento individual) e o reconhecimento associativo (ou reconhecimento de pares) são tipicamente usados na investigação desse tipo de memória. Normalmente, essas tarefas são divididas em duas etapas. Na primeira etapa (fase de codificação ou de estudo dos estímulos), os sujeitos entram em contato com os estímulos que devem ser codificados e armazenados até a segunda etapa da tarefa, que é o teste de reconhecimento. O tempo entre a primeira e a segunda fase varia de acordo com o objetivo do experimento (Anderson, 2009; Yonelinas, 2002).

No caso do reconhecimento de itens, na primeira fase, deve-se codificar uma lista de estímulos apresentados individualmente e durante o teste de reconhecimento propriamente dito, deve-se reconhecer os mesmos estímulos apresentados anteriormente. No que diz respeito ao reconhecimento associativo, na fase de codificação, são apresentados pares de estímulos, por exemplo, A-B, C-D e E-F. Na segunda etapa, são apresentados os pares de estímulos estudados previamente (pares intactos, A-B) e pares de estímulos reagrupados, que seriam, por exemplo, os pares A-C e B-D, que contém elementos que pertenciam a diferentes pares na fase de codificação (Anderson, 2009; Yonelinas, 2002).

Desse modo, em um teste de reconhecimento de itens os participantes devem discriminar entre os estímulos estudados e os estímulos novos e para essa discriminação, o processo de familiaridade é suficiente (Yonelinas, 2002; Yonelinas & Jacoby, 1994). No caso do reconhecimento de pares, os participantes codificam na lista de estudo, pares de estímulos, como A-B e C-D. Posteriormente, devem discriminar entre pares intactos e pares reagrupados e apenas a atuação do processo de familiaridade não é suficiente, porque todos os estímulos que compõe os pares são familiares, já que todos foram estudados na lista de codificação. Nesse caso, é necessária a atuação do processo de lembrança, ou seja, o participante precisa se lembrar de que o estímulo A foi estudado no mesmo par que o estímulo B e não C (Donaldson & Rugg, 1999; Yonelinas, 1997).

Há, ainda, evidências de estudos de potencial de evento relacionado (ERP) e de neuroimagem funcional que apoiam a separação dos processos de familiaridade e de lembrança e indicam que cada um dos processos se relaciona com diferentes relatos eletrofisiológicos e regiões cerebrais (Aggleton & Brown, 2006; Rugg & Curran, 2007; Woodruff et al., 2006). Por exemplo, a familiaridade foi relacionada com um efeito eletrofisiológico frontal medial bilateral que se manifesta de 300 a 500ms após a apresentação do estímulo (Curran 2004; Woodruff et al.,

2006). Esse componente é rápido e precede as tentativas controladas e conscientes do indivíduo para se lembrar da informação. Além disso, o uso da familiaridade em tarefas de reconhecimento de itens foi associado com a ativação do córtex perirrinal (Aggleton & Brown, 2006; Eichenbaum, Yonelinas, & Ranganath, 2007; Rugg & Curran, 2007; Woodruff et al., 2006).

O processo de lembrança subjacente ao reconhecimento foi associado a um efeito eletrofisiológico localizado nos eletrodos das regiões parietais, principalmente esquerda, que tem inicio entre 500 e 800ms após a apresentação do estímulo, ou seja, é iniciado depois do componente relacionado com a familiaridade (Curran, 2004; Rugg et al., 1998; Woodruff et al., 2006). Esse componente também acompanha as respostas de reconhecimento associativo ou de pares (Donaldson & Rugg, 1999). Em relação às áreas cerebrais diferencialmente ativadas durante as tarefas de reconhecimento de pares (baseadas na lembrança), foi observada a ativação do hipocampo e o parahipocampo posterior (Eichenbaum et al., 2007).

Portanto, existe uma tendência em aceitar a participação da familiaridade e da lembrança como base para as respostas de reconhecimento. Foi observado que em tarefas de reconhecimento de item, a familiaridade é o principal componente do reconhecimento e as respostas dadas nesse tipo de teste são acompanhadas por um efeito eletrofisiológico característico e pela ativação diferencial do córtex perirrinal. As respostas dadas durante as tarefas de reconhecimento associativo são baseadas na lembrança e são acompanhadas pelo efeito eletrofisiológico típico da lembrança e da ativação diferencial do hipocampo e do córtex parahipocampal (Aggleton & Brown, 2006; Eichenbaum et al., 2007; Rugg & Curran, 2007).

Além da estrutura dos itens, a memória de reconhecimento também é influenciada pelo tipo de estímulo utilizado, por exemplo, se são figuras ou palavras ou se há correspondência perceptiva entre os estímulos apresentados na fase de codificação e os apresentados na fase de reconhecimento (Mintzer & Snodgrass, 1999; Schloerscheidt & Rugg, 2004; Stenberg, 2006; Stenberg, Radeborg, & Hedman, 1995). Uma condição de correspondência perceptiva é quando uma palavra é apresentada na fase de codificação e, posteriormente, a mesma palavra é apresentada no teste de reconhecimento. Quando não há correspondência perceptiva, uma palavra aparece na fase de codificação, mas na fase de reconhecimento, o participante entra em contato com uma figura (ou vice-versa). Outra possível condição é a apresentação, na fase de codificação, de uma figura que representa um objeto ou ser vivo (por exemplo, um cachorro) e, no teste de reconhecimento, é apresentada uma figura diferente que representa o mesmo objeto ou ser vivo (Mintzer & Snodgrass, 1999; Stenberg et al., 1995).

A presença de figuras ou palavras durante a codificação é uma condição que afeta a taxa de acerto e erro em tarefas de reconhecimento. Nesse contexto, durante décadas foi observado o denominado efeito da *superioridade das figuras*, que se entende pelo reconhecimento mais preciso

e mais rápido após a codificação de figuras em comparação com a codificação das palavras, como observado por Mintzer e Snodgrass (1999), Schloerscheidt e Rugg (2004), Stenberg (2006) e Stenberg et al. (1995). Nesses estudos, foi observado que se há correspondência perceptiva entre os estímulos apresentados na codificação e no teste (FF e PP), o reconhecimento é superior em comparação com a mudança perceptiva entre lista de codificação e teste (FP e PF). Além disso, devido ao efeito da superioridade das figuras, quando se codificam e reconhecem figuras (FF), os itens são mais fáceis que quando se codificam e reconhecem palavra (PP). Assim, os estudos de Mintzer e Snodgrass (1999), Schloerscheidt e Rugg (2004) e Stenberg et al. (1995), mostraram que o padrão de respostas desses itens foi que a condição de figura-figura (FF) obteve o melhor rendimento, enquanto que as condições de figura-palavra (FP) e de palavra-palavra (PP) obtiveram um número de acerto praticamente igual e, finalmente, a condição de palavra-figura (PF) foi a mais difícil.

Ao lado disso, quando essas condições experimentais foram aplicadas em um paradigma de reconhecimento associativo por Hockley (2008), foi observado um padrão de respostas muito similar. O autor levou a cabo quatro experimentos nos quais os participantes (estudantes universitários) deviam estudar uma lista de 60 pares de figuras e palavras que durante o teste de reconhecimento eram apresentados como pares intactos ou reagrupados. Em todos os experimentos, se manteve a superioridade do reconhecimento das figuras em relação ao reconhecimento de palavras. No quarto experimento, as figuras estudadas foram reconhecidas como palavras e ainda assim foram reconhecidas com mais facilidade que as palavras codificadas, mas não tão bem quanto os itens de codificação e reconhecimento de figuras. A taxa de acerto foi superior para os pares estudados e reconhecidos como figuras (FF), os pares estudados como figuras e reconhecidos como palavras (FP) foram mais fáceis que os pares modais de palavras (PP) e os pares de palavras e figuras (PF) foram os mais difíceis. Portanto, observou-se, também no caso dos pares, o efeito da superioridade pictórica e o aumento da dificuldade produzido pela mudança de formato dos estímulos entre codificação e teste.

Finalmente, a memória de reconhecimento também varia em função das características dos sujeitos. Nesse sentido, as mulheres apresentam, em média, desempenhos superiores aos dos homens em tarefas de memória de reconhecimento visual (Halpern, 2012; Kimura, 1999; McGivern et al., 1997; McGivern et al., 1998). Por isso, espera-se que em um teste de memória de reconhecimento visual, as mulheres apresentem um desempenho superior ao dos homens.

Assim, nos últimos anos, tem-se observado uma tendência na literatura em considerar que existem dois processos subjacentes ao reconhecimento. Por isso, a avaliação desse processo de memória deve ser mais precisa e levar em consideração o tipo de tarefa que avalia melhor cada um dos processos, como uma tarefa de reconhecimento de

itens para avaliar, principalmente, a familiaridade e uma tarefa de reconhecimento associativo para a medição do uso da lembrança como base para o reconhecimento. Não é de conhecimento dos autores a existência, no Brasil, de um teste de memória de reconhecimento que faça uma diferenciação explícita entre a participação relativa de cada processo em função do tipo de tarefa. Assim, é necessário desenvolver instrumentos de medida que levem em consideração os avanços teóricos e as discussões atuais relacionadas com a memória de reconhecimento. Além disso, são necessários novos testes que possibilitem uma maior padronização da apresentação de estímulos e uma melhoria do conteúdo dos itens. Por isso, este estudo teve o objetivo de desenvolver uma versão computadorizada para um teste de reconhecimento associativo, assim como desenvolver itens para a construção e informatização de um teste de reconhecimento de itens, utilizando o modelo de Rasch (1960). Também foi objetivo deste estudo testar as hipóteses dos níveis de dificuldade dos itens, conforme previsto pelos experimentos de correspondência perceptiva, além de buscar evidência de validade para os instrumentos, averiguando se as mulheres obtêm desempenho superior ao dos homens.

## Método

### Participantes

Participaram da pesquisa 217 estudantes de uma universidade particular do interior do estado de São Paulo, dos quais aproximadamente 50% eram alunos de psicologia, 40% de engenharia e 9,7% eram estudantes de matemática. Além disso, 53,5% eram do sexo masculino. A média de idade foi de 25 anos ( $DP=6,67$ ). Todos os participantes eram voluntários e assinaram o Termo de Consentimento Livre e Esclarecido.

### Instrumento

Foram aplicados dois testes computadorizados, quais sejam, um teste de memória de reconhecimento associativo e um teste de reconhecimento de itens. Ambos os testes de reconhecimento foram computadorizados com o uso do programa *Revolution 3.0 Studio Edition*. O teste de reconhecimento associativo se baseou em uma versão prévia construída para aplicações coletivas (Fernandes, Prieto, & Delgado, 2010). O desenvolvimento dos itens dessa primeira versão se baseou nos mesmos estudos de mudança de formato de Mintzer e Snodgrass (1999) e Stenberg et al. (1995). Levou em consideração as mesmas condições estímulares, ou seja, membros dos pares pertencentes a mesma categoria ou a categorias diferentes, codificação pictórica e de palavras e a mudança ou manutenção do formato dos estímulos entre a fase de estudo e de teste. Nessa versão, não havia a categoria de codificação de figura e reconhecimento de figuras diferentes. Para a coleta de dados, os estímulos da fase de codificação foram apresentados, em sala de aula e coletivamente, com *data show*. Esse teste também mostrou propriedades psicométricas adequadas (Fernandes et al., 2010).

Os dois testes são autoaplicativos e todas as instruções aparecem no próprio programa. Primeiramente, são iniciadas as instruções do teste de reconhecimento de itens, com explicações verbais e pictóricas, que mostram todas as condições de estímulos que serão apresentadas no teste. Após isso, é passado um exemplo de uma lista de codificação que reproduz exatamente as condições do teste e finalmente, apresenta-se um exemplo de um teste de reconhecimento. Finalmente, é apresentada a lista de codificação com os estímulos que o participante deve aprender, seguida do teste de reconhecimento. Posteriormente, são exibidas as explicações sobre o teste de reconhecimento de pares, que também são pictóricas e verbais e representam todas as condições que são encontradas no teste. É apresentado um exemplo de uma lista de codificação de pares de estímulos e um exemplo de teste de reconhecimento de pares. Quando os participantes terminam os exemplos, iniciam a fase de codificação dos pares de estímulos, seguida do teste de reconhecimento associativo.

Em relação aos itens dos testes, o teste de reconhecimento de itens inclui uma lista de codificação de 32 estímulos e um teste de reconhecimento de 64 itens, sendo 32 itens compostos por estímulos novos (distratores). O teste de pares inclui uma lista de codificação de 40 pares e um teste de reconhecimento de 64 itens, sendo 24 pares

reagrupados. Em ambos os testes, cada item de reconhecimento possui seis opções de respostas, sendo que a última opção indica que o estímulo ou o par estudado não está presente. O tempo de codificação de cada estímulo ou par de estímulos é de dois segundos.

Os diferentes tipos de itens do teste de reconhecimento de itens foram criados e agrupados em categorias, que são: figura-figura (I-FF), figura-figura diferente (I-FD), figura-palavra (I-FP), palavra-palavra (I-PP), palavra-figura (I-PF). Em relação aos distratores, foi apresentado ao longo do teste um total de 72. Cada um deles apareceu quatro vezes (em quatro itens diferentes) para completar todas as opções de resposta necessárias (288). Os distratores também foram apresentados no mesmo formato ou em formato diferentes (figuras ou palavras), então 18 foram apresentados somente como figuras, 18 somente como palavras, 18 foram apresentados duas vezes como figuras e duas vezes como palavras e 18 foram apresentados duas vezes como palavras e duas como figuras. Segue um exemplo de um item de reconhecimento individual nas Figuras 1 e 2. A Figura 1 representa um estímulo de codificação no teste de itens (palavra escrita *unha*) e a Figura 2 representa um item apresentado no teste de reconhecimento individual, no qual a resposta correta seria a própria palavra escrita *unha*.

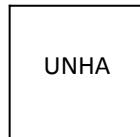


Figura 1. Item de codificação do teste de reconhecimento individual.



Figura 2. Item de reconhecimento do teste individual.

O teste de reconhecimento associativo inclui 10 categorias de pares. Em cinco categorias, os elementos dos pares pertencem a mesma categoria semântica, como meio de transporte, animais, frutas, etc. Nas outras cinco categorias, os elementos dos pares pertencem a categorias diferentes. As categorias são: figura-figura associada (P-FFA), figura-figura não associada (P-FFN), figura-figura diferente associada (P-FDA), figura-figura diferente não associada (P-FDN), figura-palavra associada (P-FPA), figura-palavra não associada (P-FPN), palavra-palavra associada (P-PPA), palavra-palavra não associada (P-PPN), palavra-figura associada (P-PFA), palavra-figura não associada (P-PFN). Os 24 pares reagrupados foram

apresentados da seguinte forma: seis foram apresentados no formato FF e FD (três em cada formato) e seis pares em cada um dos formatos restantes (FP, PP, PF). Os estímulos de *recheio* são aqueles que aparecem nas opções de resposta dos itens, que não são nem os estímulos estudados nem a última opção de resposta (o par não está presente). Em cada item aparece sempre um par de estímulos estudado ou agrupado, a última opção que indica que o par não está presente e, por isso, cada item sempre possui três opções que devem ser completadas com outros estímulos. Na Figura 3, observa-se um exemplo de um par de estímulo de codificação e na Figura 4 está representado um item de reconhecimento associativo.

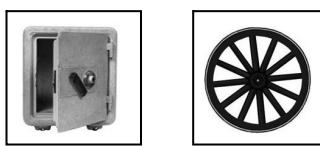


Figura 3. Par de codificação do teste de reconhecimento associativo.

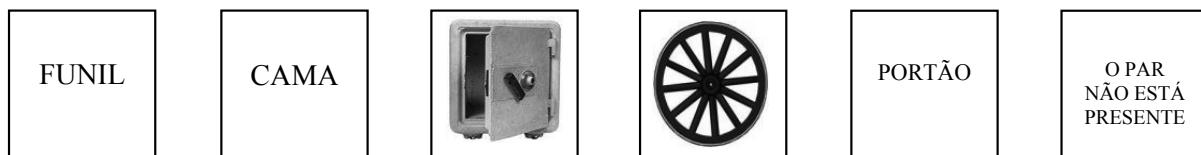


Figura 4. Par de reconhecimento do teste de reconhecimento associativo.

### Procedimento

O projeto foi inicialmente aprovado pelo Comitê de Ética em pesquisa da universidade onde foram coletados os dados. Os testes foram aplicados nas salas de informática da universidade. Com o consentimento dos professores, os alunos se deslocaram às salas, nas quais se havia instalado previamente o programa que continha o teste. Como o teste é auto-administrativo, as instruções estavam no programa e os participantes davam inicio aos dois testes. Na sala sempre havia um aplicador para resolver eventuais dúvidas e complicações. As aplicações de ambos os testes duraram por volta de 25 minutos.

### Procedimento de Análise dos Dados

A análise psicométrica dos instrumentos foi realizada pelo modelo de Rasch (1960), explicado a seguir, e com o software WINSTEPS, versão 3.69.0 (Linacre, 2009). A análise estatística de comparação entre o desempenho masculino e o feminino (medidas obtidas pela análise de Rasch) foi realizada pela prova *t* de Student e a comparação entre médias de dificuldade das categorias dos itens, tanto individuais como de pares (também obtidas pela análise de Rasch), foi realizada pela ANOVA, com o uso do software SPSS-IBM.

O modelo de Rasch (1960) é um modelo probabilístico de um parâmetro que estima o nível de dificuldade dos itens e de habilidade dos sujeitos separada e independentemente. A fórmula da probabilidade de acerto de um sujeito para cada item estima conjuntamente os parâmetros de dificuldade dos itens e habilidade dos sujeitos e os únicos valores que afetam a probabilidade de acerto são a habilidade do sujeito e a dificuldade do item. O modelo de Rasch estima simultaneamente os parâmetros das pessoas e dos itens, por meio de métodos de máxima verossimilhança. Essa medição conjunta é realizada de forma independente e, por isso, é possível generalizar as estimativas da dificuldade dos itens de forma independente da amostra utilizada ao mesmo tempo em que permite generalizar as estimativas dos sujeitos independentemente dos itens respondidos (Prieto & Delgado, 2003).

A utilização do modelo de Rasch (1960) requer que os dados obtidos empiricamente se ajustem ao modelo, ou

seja, que os sujeitos respondam conforme a expectativa do modelo. Esse ajuste ou desajuste dos dados é obtido pelos índices de *infit* e *outfit* dos sujeitos e dos itens e revelam se há uma maior ou menor variação entre o padrão observado de respostas dos sujeitos e o padrão esperado de respostas. Linacre (2002) considera que os valores entre 0,5 e 1,5 revelam divergências pouco relevantes. Os valores de 1,5 a 2,0, a sua vez, indicam que as divergências são moderadas, enquanto que os valores superiores a 2,0 são considerados aberrantes. Valores inferiores a 0,5 manifestam maior ajuste do esperado, mas ao acaso e não chegam a ser degradantes.

No que diz respeito aos estatísticos de fidedignidade dos itens e dos sujeitos, denominados *Person Separation Reliability* e *Item Separation Reliability*, seus valores, que podem variar entre 0 e 1, indicam se a localização dos itens e das pessoas na escala da variável latente é replicável (Linacre, 2009). A fidedignidade dos sujeitos é análoga ao alfa de Cronbach e o KR-20.

Outra análise realizada a partir do modelo é a do *funcionamento diferencial dos itens* (DIF), que busca estabelecer se há viés nos itens em relação a diferentes grupos, como homens e mulheres. Espera-se que pessoas de diferentes grupos, mas que possuam o mesmo nível de habilidade, tenham a mesma probabilidade de acerto ao item. Se a probabilidade de acerto varia consideravelmente para pessoas de diferentes grupos com o mesmo nível de habilidade, considera-se a possibilidade de que outra variável ou outra habilidade esteja exercendo influência na probabilidade de resposta ao item. Nesse caso, há um problema de dimensionalidade do item que faz com que o item seja mais fácil para um grupo, em detrimento de outro, mesmo que tenham habilidades semelhantes.

Os critérios para a detecção do DIF são o tamanho da diferença entre os estimadores da dificuldade nos grupos (homens e mulheres) e a sua significância. Considera-se que diferenças iguais ou superiores a 0,50 *logits* são relevantes. Neste estudo, valores positivos de contraste favorecem aos homens e valores negativos beneficiam às mulheres. Para a significância estatística, Linacre (2009) sugeriu fazer a correção de Bonferroni, na qual o valor de  $p=0,05$  é dividido pelo número total de comparações (32

itens no teste individual e 40 no teste de pares). Por isso, os valores de  $p$  inferiores a 0,00156 (teste individual) e a 0,00125 (teste de pares) indicam que o DIF é significativo e não se deve ao acaso. Nesta pesquisa, foi realizada a análise de DIF associado ao sexo. Como um dos objetivos do estudo foi buscar diferenças de desempenho entre homens e mulheres, a ausência de DIF permite afirmar que a diferença entre homens e mulheres não é artificial, mas se deve a uma diferença válida entre os níveis de habilidade dos homens e mulheres.

## Resultados

Inicialmente, são apresentados os resultados relativos à análise psicométrica pelo modelo de Rasch (1960) e posteriormente as análises estatísticas de diferença entre médias de dificuldade de categoria de itens e entre médias

de habilidades de homens e mulheres. Para a análise dos itens, foram considerados apenas os itens estudados e não os distratores. Então foram analisados 32 itens individuais e 40 pares de estímulos.

### Teste de Reconhecimento de Itens

A estimativa dos parâmetros dos itens e das pessoas produziu o mapa de Wright apresentado na Figura 5. Observa-se que os itens e as pessoas se distribuem adequadamente ao longo da escala, com uma maior concentração de itens entre 0,5 e 1,5 *logits*. A maioria de sujeitos está dentro do intervalo medido pelos itens do teste, o que indica que esses itens são adequados para medir as habilidades dessas pessoas. Também pode ser observado no mapa que são manifestadas as diferenças individuais entre os sujeitos, dada a sua distribuição no continuo.

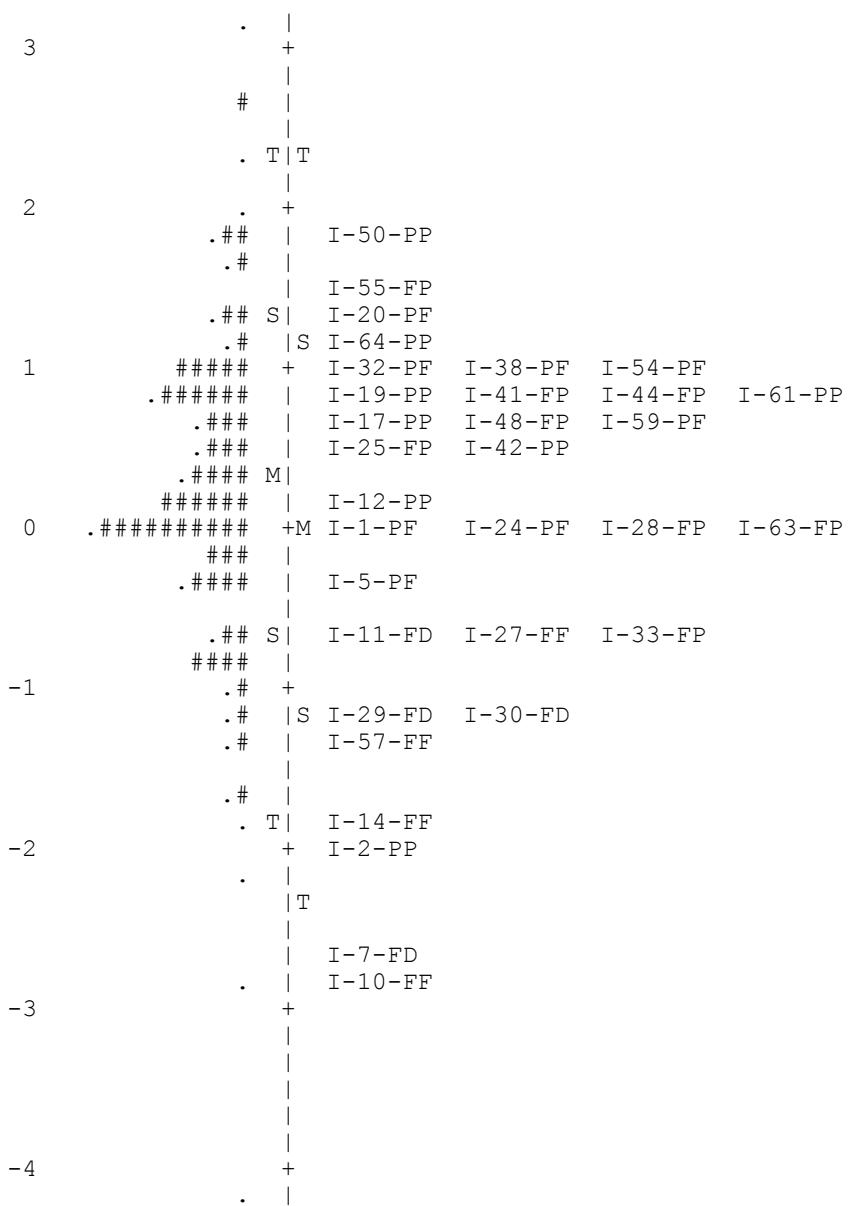


Figura 5. Mapa do construto para o teste individual.

Nota. Cada “#” são 3. Cada “.” é 1 ou 2.

A média da habilidade dos sujeitos foi 0,26 *logits* ( $DP=1,01$ ) e os valores mínimos e máximos foram de -4,20 a 3,17 *logits*, respectivamente, o que indica que suas habilidades oscilaram amplamente. O erro de medida das pessoas foi relativamente baixo (média de 0,43 *logits* e  $DP=0,006$ ) e a confiabilidade de suas pontuações foi 0,80, medida tanto pelo alfa de Cronbach quanto pelo *Person Separation Reliability*.

Ainda em relação às análises dos sujeitos, obteve-se uma diferença significativa entre os homens e as mulheres. A média de habilidade das mulheres foi de 0,46 *logits*, enquanto a média de habilidade dos homens foi de 0,09 *logits*. A comparação entre as médias mediante a prova *t* de Student indicou que essa diferença de 0,37 *logits* é

significativa ( $p=0,006$ ), o que mostra que a média das mulheres foi significativamente superior à dos homens no teste de reconhecimento de itens.

No que diz respeito às análises dos itens, a média da dificuldade dos itens foi, por convenção, estabelecida como zero, com valores mínimos e máximos de dificuldade de -2,83 a 1,79 *logits*, respectivamente. O erro de medida dos itens foi baixo (média de 0,17 e  $DP=0,03$ ). A precisão dos itens, medida pelo *Item Separation Reliability*, foi de 0,98, um valor muito alto. No que diz respeito ao ajuste, a média de *infit* foi de 0,99 ( $DP=0,06$ ) e de *outfit* foi de 1,05 ( $DP=0,22$ ).

A Tabela 1 mostra a média de dificuldade e o desvio padrão das categorias de itens elaboradas. Observa-se que as categorias seguiram a ordem de dificuldade prevista.

Tabela 1  
Média e Desvio Padrão da Dificuldade das Categorias

Categorias	Média	Desvio Padrão
Individual figura-figura (I-FF)	-1,66	0,89
Individual figura diferente (I-FD)	-1,45	0,89
Individual figura-palavra (I-FP)	0,48	0,65
Individual palavra-palavra (I-PP)	0,48	1,12
Individual palavra-figura (I-PF)	0,60	0,62

Para averiguar se as diferenças entre as médias das categorias são significativas, foi levado a cabo uma análise de variância (ANOVA) e a prova a posteriori de Bonferroni. A ANOVA indicou que as diferenças entre as médias não se devem ao acaso [ $F(4,217)=9,074, p<0,01$ ]. A prova de Bonferroni mostrou que não há diferença significativa entre as categorias de I-FF e a de I-FD, mas sim entre a categoria de I-FF e as categorias de I-FP, I-PP e I-PF. A média da categoria de I-FD se diferenciou significativamente das categorias I-FP, I-PP e I-PF, enquanto que a categoria I-FP se diferenciou significativamente das médias das categorias I-FF e I-FD. A média de PP apresentou diferença significativa em relação a I-FF e I-FD e a média de I-PF se diferenciou de forma significativa de I-FF e I-FD. São formadas, portanto, duas agrupações com níveis de dificuldades significativamente distintos: por um lado, estão as categorias I-FF e I-FD, que são as mais fáceis e, por outro lado, estão as categorias I-FP, I-PP e I-PF, cujas médias são mais difíceis.

Posteriormente, para averiguar se a ordem de dificuldade proposta das categorias corroborava com a ordem obtida

empiricamente foi realizada a análise de correlação ordinal de Spearman; a correlação foi de 0,975 ( $p=0,005$ ), o que indica que as categorias apresentaram a ordem crescente de dificuldade esperada e que, como havia sido previsto, as categorias I-FP e I-PP obtiveram médias equivalentes.

Para a análise do DIF, foram considerados os critérios de Linacre (2009). Dos 32 itens analisados, 15 apresentaram um contraste de tamanho considerável, mas apenas um deles, o item I-17-PP foi significativo. Esse item apresentou um contraste de 1,18 e valor de  $p=0,0002$ , beneficiando diferencialmente os homens. Portanto, apenas um item apresentou viés em relação aos gêneros.

#### Teste de Reconhecimento Associativo

O mapa do construto, exibido na Figura 6, indica que há uma considerável distribuição dos sujeitos e dos itens ao longo da escala e que a maioria deles se localiza dentro do intervalo aproximado de -1 a +1 *logit*. Ademais, a maioria de sujeitos se encontra dentro do intervalo medido pelos itens, o que mostra que os pares são adequados para medir essa amostra.

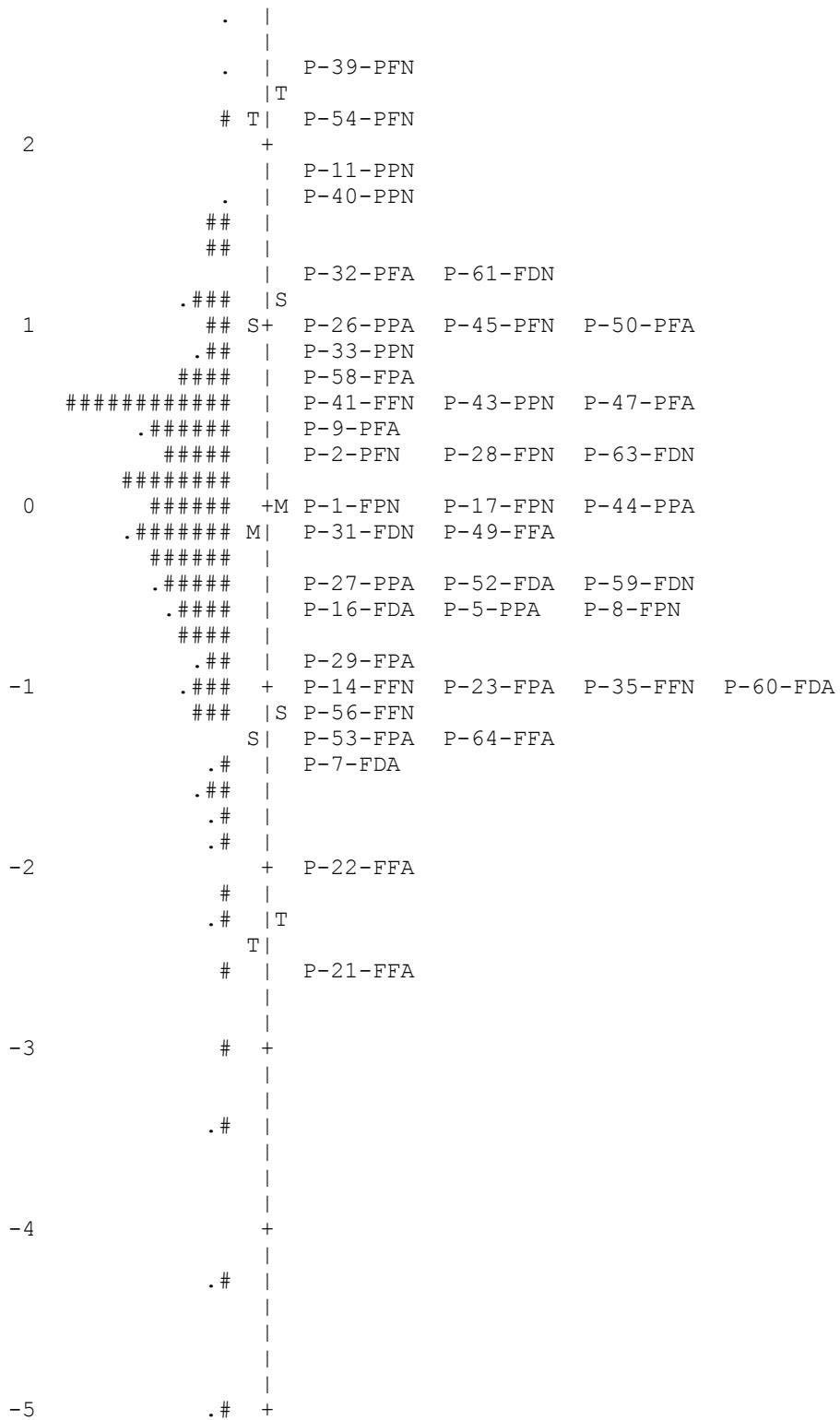


Figura 6. Mapa do construto para o teste de pares.

Nota. Cada “#” são 2. Cada “.” é 1.

Em relação aos estatísticos das pessoas no teste de pares, a média da habilidade das pessoas foi  $-0,18$  *logits* ( $DP=1,15$ ), com valores mínimos e máximos de  $-4,22$  a  $2,65$ , respectivamente. O valor do alfa de Cronbach foi de  $0,88$  e o da confiabilidade medida pelo *Person Separation*

*Reliability* foi de  $0,87$ . O erro padrão das pessoas foi em média  $0,39$  ( $DP=0,10$ ). Em relação ao ajuste, os sujeitos apresentaram uma média de *infit* de  $1,00$  ( $DP=0,15$ ) e de *outfit* de  $1,13$  ( $DP=0,84$ ). Novamente, a comparação entre homens e mulheres revelou que as mulheres obtiveram um

rendimento superior aos homens. A prova *t* mostrou que a média das mulheres foi de 0,25 *logit* e a dos homens de -0,69 *logits* e que essa diferença de 0,94 *logit* foi significativa ( $p<0,01$ ).

No que diz respeito à análise dos pares, a média de dificuldade foi estabelecida como zero ( $DP=1,11$ ) e os valores mínimos e máximos obtidos foram -2,61 e 2,38 *logits*. O erro padrão dos itens foi em média 0,17 ( $DP=0,2$ )

e a precisão dos itens, medida pelo índice *Item Separation Reliability*, foi de 0,98. Os itens novamente se ajustaram ao modelo; a média do *infit* foi de 0,99 ( $DP=0,08$ ), e a média do *outfit* foi 1,13 ( $DP=0,54$ ).

Em relação às 10 categorias de dificuldade propostas, não foi obtida exatamente a ordem de dificuldade esperada. A Tabela 2 mostra os valores das médias das categorias na ordem crescente de dificuldade encontrada e a ordem esperada.

Tabela 2  
*Média de Dificuldade das Categorias de Itens e Ordem Proposta de Dificuldade*

Categorias	Ordem proposta	Médias de dificuldade	DP
Par figuras-figuras associadas (P-FFA)	1	-1,54	1,05
Par figuras-figuras diferentes associadas (P-FDA)	3	-0,88	0,43
Par figuras-figuras não associadas (P-FFN)	2	-0,64	0,80
Par figuras-palavras associadas (P-FPA)	5	-0,62	0,88
Par figuras-palavras não associadas (P-FPN)	6	-0,05	0,36
Par palavras-palavras associadas (P-PPA)	7	0,01	0,73
Par figuras-figuras diferentes não associadas (P-FDN)	4	0,24	0,75
Par palavras-figuras associadas (P-PFA)	9	0,84	0,39
Par palavras-palavras não associadas (P-PPN)	8	1,22	0,68
Par palavras-figuras não associadas (P-PFN)	10	1,42	0,99

A correlação ordinal de Spearman entre a ordem de dificuldade proposta e a ordem de dificuldade obtida revelou um valor de 0,915 ( $p<0,01$ ), o que indica uma alta correlação entre a ordem proposta e a obtida empiricamente. A ANOVA das médias das 10 categorias indicou que [ $F(9,217)=6,580$ ,  $p<0,01$ ], sugerindo que as diferenças entre as médias não podem ser atribuídas ao acaso. A prova a posteriori de Bonferroni indicou que a categoria Par figuras-figuras associadas se diferenciou significativamente das categorias Par palavras-palavras não associadas, Par palavras-figuras associadas e Par palavras-figuras não associadas, mas não das categorias Par figuras-figuras não associadas, Par figuras-figuras diferentes associadas, Par figuras-figuras diferentes não associadas, Par figuras-palavras associadas, Par palavras-palavras associadas e Par figuras-palavras não associadas. A categoria Par figuras-figuras não associadas não foi significativamente diferente de nenhuma categoria, exceto a de Par palavras-figuras não associadas, enquanto que a categoria Par figuras-figuras diferentes associadas apresentou diferença significativa das categorias par palavras-palavras não associadas e Par palavras-figuras não associadas.

A categoria Par figuras-figuras diferentes não associadas não se distinguiu de nenhuma categoria, mas a de Par figuras-palavras associadas se diferenciou de forma significativa da categoria de Par palavras-palavras não associadas e de Par palavras-figuras não associadas. A categoria Par palavras-palavras associadas não se diferenciou das demais, assim como a de par figuras-palavras não associadas. Não obstante, as categorias Par palavras-palavras não associadas e Par palavras-figuras associadas foram significativamente diferente da Par figuras-figuras associadas. Finalmente, a categoria Par palavras-figuras não associadas foram significativamente diferentes das categorias Par figuras-figuras associadas, Par figuras-figuras não associadas, Par figuras-figuras diferentes associadas e Par figuras-palavras associadas.

Em relação ao estudo do funcionamento diferencial dos itens, novamente foram considerados os critérios de Linacre (2009). A correção de Bonferroni resultou em um nível de significância de  $p=0,00125$ . Do total de pares, onze apresentaram um contraste considerável (maior que 0,50 *logits*), mas somente o contraste do par P-60-FDA, de 1,1 *logit*, foi significativo e favoreceu os homens.

## Discussão

Os principais objetivos desta pesquisa foram desenvolver uma versão computadorizada para um teste de reconhecimento associativo, assim como desenvolver itens para a construção e informatização de um teste de reconhecimento de itens. Além disso, foram analisadas as pontuações dos itens e sujeitos utilizando o modelo de Rasch (1960) e foram obtidas evidências sobre a validade de construto com base na correspondência entre as ordens de dificuldade das categorias de itens individuais e de pares de estímulos previstas e as ordens de dificuldade encontradas empiricamente. Do mesmo modo, a diferença de desempenho entre homens e mulheres obtida neste estudo, que não se deve ao DIF, está de acordo com a literatura (Halpern, 2012; Kimura, 1999) e, por isso, é uma evidência de validade de construto.

Para ambos os testes, a análise de Rasch (1960) baseada nos indicadores *infit* e *outfit* indicou que a maioria dos itens se ajustaram ao modelo de Rasch. Esse ajuste significa que os itens foram respondidos pelos participantes segundo o esperado. Os itens apresentaram valores adequados de fidedignidade e, portanto, a ordem de dificuldade dos itens é estável e será mantida em aplicações futuras (Linacre, 2009). Houve uma distribuição considerável dos itens ao longo do continuo da memória de reconhecimento, o que mostra que há itens com diferentes níveis de dificuldade, tanto itens mais fáceis como mais difíceis.

Em relação ao funcionamento diferencial dos itens, apenas um item de cada teste apresentou DIF. Isso quer dizer que mesmo que homens e mulheres apresentem o mesmo nível de habilidade na memória de reconhecimento, o item será mais difícil para as mulheres (Linacre, 2009). Esses itens devem ser avaliados, modificados ou eliminados, pois há um problema de dimensionalidade, ou seja, deve haver alguma variável externa que está influenciando a probabilidade de resposta desses itens. Não obstante, esse é o caso apenas de dois itens, e por isso, pode-se afirmar que os testes não apresentam viés em relação aos homens e mulheres. Essa constatação é importante, pois garante que as diferenças de capacidade mnemônica encontradas entre homens e mulheres não se deve aos itens dos testes e sim às características dos sujeitos.

No que diz respeito às pessoas, a fidedignidade de suas pontuações foi alta e, por isso, a ordem de habilidade também é estável e se houvessem aplicações futuras, ela seria mantida com considerável precisão (Linacre, 2009). Igualmente, houve uma distribuição apropriada ao longo da escala, permitindo uma adequada representação das diferenças individuais dos sujeitos. Foi obtido também um bom ajuste, o que mostra que as pessoas, de forma geral, responderam segundo a expectativa do modelo. Isso indica que pessoas com boa memória de reconhecimento acertaram itens fáceis, médios, inclusive mais difíceis e que houve também coerência em seus erros, quer dizer, não erraram itens fáceis, apenas itens mais difíceis. Portanto, esse conjunto de resultados indica que

ambos os testes apresentam propriedades psicométricas adequadas.

No que diz respeito às categorias de dificuldade dos itens, construídas em função de experimentos realizados anteriormente sobre ausência e presença de correspondência perceptiva (Mintzer & Snodgrass, 1999; Schloerscheidt & Rugg, 2004; Stenberg et al., 1995), observou-se no caso do teste de reconhecimento individual o mesmo padrão de dificuldade, corroborando com os resultados desses experimentos. Foi, portanto, observado o efeito da superioridade das figuras e a maior facilidade de reconhecimento de palavras e figuras quando não há mudança de formato entre os estímulos apresentados na fase de codificação e no teste. Nesse sentido, a ordem de dificuldade obtida reflete a interação esperada entre as variáveis introduzidas nos itens. A correspondência entre a dificuldade empírica e teórica das condições com as quais foram gerados os itens é a principal evidência de construto encontrada neste estudo.

Os dados mostraram duas agrupações de dificuldades dos itens que são significativamente distintas. É interessante notar que as categorias de codificação e reconhecimento de figuras e de codificação e reconhecimento de figuras diferentes foram muito fáceis e não se diferenciaram significativamente. Isso significa que uma pessoa reconhece com a mesma facilidade a figura de um cachorro vista anteriormente ou a figura de um cachorro diferente. Assim, em termos de avaliação da capacidade de memória, esses itens estão dentro de um mesmo intervalo de dificuldade e não discriminam entre sujeitos com diferentes níveis de habilidade. Por isso, não está clara a relevância da mudança perceptiva neste caso para avaliar a memória de reconhecimento.

As outras três categorias de itens não se diferenciaram entre si em termos de dificuldade de acerto, mas esse agrupamento pode ser compreendido pelos tipos de estímulos. Mesmo que esse agrupamento tenha uma categoria de itens de codificação de figuras, o reconhecimento desses itens é mais prejudicado quando há mudança de formato entre codificação e teste (Mintzer & Snodgrass, 1999; Schloerscheidt & Rugg, 2004; Stenberg et al., 1995) e por isso são itens relativamente mais difíceis. No caso da codificação e reconhecimento de palavras, o reconhecimento desses itens não tem o benefício da codificação pictórica e da consequente superioridade das figuras. Mas o reconhecimento também não é prejudicado pela mudança de formato entre estudo e teste. E finalmente, codificar palavras e reconhecer figuras foi o tipo de item mais difícil, como era esperado (Mintzer & Snodgrass, 1999; Schloerscheidt & Rugg, 2004; Stenberg et al., 1995), no entanto, não se diferenciou das outras categorias. A codificação de palavras é menos eficiente que a de figuras e, além disso, há a desvantagem da mudança de formato entre estudo e teste, por isso esperava-se que os itens fossem significativamente mais difíceis. É interessante ressaltar que apesar de não haver diferença significativa entre as dificuldades das categorias de itens desse agrupamento, pode ser observado no mapa de Wright que eles se distribuem bem nos níveis mais altos

de dificuldade na escala *logit* e promovem certa variabilidade de dificuldade dentro desse intervalo.

Em termos de subprocessos do reconhecimento, apesar desse estudo não testar a relação entre a familiaridade e a tarefa de reconhecimento de itens, com base na definição do processo de familiaridade, sugere-se que as duas categorias mais fáceis de itens (figura-figura e figura-figura diferente) exigem menos esforço consciente de busca e de lembrança e que se baseiam mais na sensação de familiaridade dos estímulos, produzida rápida e automaticamente após o reencontro com o estímulo (Curran, 2004; Yonelinas, 1994, 2002). Os itens mais difíceis devem exigir mais recursos relacionados com a busca na memória dos estímulos codificados e, por isso, deve haver uma maior participação da lembrança do que nos itens mais fáceis. Ressalta-se que a literatura indica que o processo de reconhecimento depende tanto da familiaridade e da lembrança e que não há nenhum paradigma de reconhecimento que desencadeie a participação de somente um deles (Yonelinas, 1994, 1997, 2002).

Em relação à ordem de dificuldade das categorias de pares de estímulos, foi observada uma alta correlação entre a ordem esperada e a obtida empiricamente, o que indica que de modo geral, a manipulação das variáveis das condições dos itens produziu os níveis de dificuldade esperados, corroborando com os achados de Hockley (2008) que estendeu o efeito da superioridade das figuras e o prejuízo do reconhecimento como consequência da mudança de formato dos estímulos para o caso do reconhecimento associativo.

A principal previsão que não foi obtida empiricamente, pois se distanciou muito do esperado, foi a categoria de pares de figuras-figuras diferentes não associadas, que apresentou um nível de dificuldade mais alto do que era esperado. Deveria ter sido a quarta categoria mais fácil, no entanto, essa categoria foi tão difícil quanto a categoria de codificação de palavras. Sugere-se que esse resultado se deva a um problema de identificação de alguma figura apresentada na fase de codificação, que não possibilitou a compreensão da figura e inviabilizou o seu futuro reconhecimento.

No que se refere aos processos subjacentes ao reconhecimento, estudos anteriores mostraram que esse tipo de tarefa de pares depende da lembrança para a discriminação entre pares reagrupados e pares intactos (Donaldson & Rugg, 1999; Eichenbaum et al., 2007; Wheeler & Buckner, 2004). Nesse teste, o papel dos processos subjacentes está mais claro que no teste de itens individuais, quer dizer, os estímulos estudados previamente durante a lista de codificação produzem uma sensação de familiaridade nos participantes (Aggleton & Brown, 2006; Eichenbaum et al., 2007; Rugg & Curran, 2007; Woodruff et. al., 2006), mas essa sensação não é suficiente. Para discriminar entre os pares, é necessária a lembrança dos dois elementos estudados no mesmo par durante a codificação (Yonelinas, 1997).

Finalmente, encontrou-se evidência de validade de construto de ambos os testes. Observou-se que as mulheres

foram significativamente mais hábeis que os homens, corroborando com resultados de estudos anteriores (Halpern, 2012; Kimura, 1999; McGivern et al., 1997; McGivern et al., 1998). Além disso, a análise de DIF mostrou que essas diferenças não se devem às condições artificiais dos testes, nem a outro construto que não seja a memória de reconhecimento.

## Considerações Finais

Este estudo mostrou as boas propriedades psicométricas dos dois testes de memória de reconhecimento, analisadas pelo modelo de Rasch (1960), tanto para os itens como para os sujeitos, tais como valores adequados de fidedignidade, boa distribuição dos níveis de dificuldade dos itens, que, além disso, revelam as diferenças individuais das pessoas avaliadas. Também não há presença preocupante de DIF. Mostrou, ainda, que a manipulação das variáveis dos itens gerou os padrões de dificuldade dos itens esperados, com poucas exceções, o que se entende como a principal evidência de validade de construto de ambos os instrumentos. Além disso, a diferença de desempenho entre homens e mulheres, observada na literatura, é considerada também validade de construto. Desse modo, as evidências de validade e os resultados psicométricos encontrados encorajam a continuidade de pesquisas dos testes, para que no futuro possam ser usados como instrumentos de avaliação da memória de reconhecimento. Ressalta-se que os testes ainda estão em estudo e sendo reformulados para seu uso futuro.

Não obstante, ainda faltam estudos que mostrem a participação relativa de cada processo subjacente ao reconhecimento, principalmente no caso do reconhecimento de itens. Esse estudo não possibilitou testar a participação dos processos para cada tipo de teste, de forma geral, ou de categoria de itens. Também é interessante ampliar a amostra de sujeitos, aplicar os instrumentos em pessoas mais novas e em idosos, em busca de outras evidências de validade.

## Referências

- Aggleton, J. P., & Brown, M. W. (2006). Interleaving brain systems for episodic and recognition memory. *Trends in Cognitive Sciences*, 10, 455-463. doi:10.1016/j.tics.2006.08.003
- Anderson, M. C. (2009). Retrieval. In A. Baddeley, M. W. Eysenck, & M. C. Anderson, *Memory* (pp. 163-189). Hove, UK: Psychology Press.
- Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia*, 42, 1088-1106. doi:10.1016/j.neuropsychologia.2003.12.011
- Defeyter, M. A., Russo, R., & McPartlin, P. L. (2009). The picture superiority effect in recognition memory: A developmental study using the response signal procedure. *Cognitive Development*, 24, 265-273. doi:10.1016/j.cogdev.2009.05.002
- Donaldson, D. I., & Rugg, M. D. (1999). Event-related potential studies of associative recognition and recall: Electrophysiological evidence for context dependent retrieval process.

- Cognitive Brain Research*, 8, 1-16. doi:10.1016/S0926-6410(98)00051-2
- Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, 30, 123-152. doi:10.1146/annurev.neuro.30.051606.094328
- Fernandes, D. C., Prieto, G. A., & Delgado, A. R. (2010). Construção de um teste transmodal de memória de reconhecimento. *Estudos de Psicologia (Campinas)*, 27, 283-291. doi:10.1590/S0103-166X2010000300001
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 91, 1-65.
- Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4<sup>th</sup> ed.). New York: Psychology Press.
- Hintzman, D. L. (1986). Schema abstraction in a multiple-trace memory model. *Psychological Review*, 93, 429-445.
- Hockley, W. E. (2008). The picture superiority effect in associative recognition. *Memory & Cognition*, 36, 1351-1359. doi:10.3758/MC.36.7.1351
- Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: The Massachusetts Institute of Technology Press.
- Linacre, J. M. (2002). What do infit and outfit, mean-squared and standardized mean? *Rasch Measurement Transactions*, 16(2), 878. Retrieved July 10, 2006, from <http://209.238.26.90/rmt/rmt82a.htm>
- Linacre, J. M. (2009). *A user's guide to WINSTEPS & MINISTEPS: Rasch Model Computers Programs*. Chicago, IL: Winsteps.com.
- Mandler, G. (2008). Familiarity breeds attempts. A critical review of dual-process theories of recognition. *Perspectives on Psychological Science*, 3, 390-399. doi:10.1111/j.1745-6924.2008.00087.x
- McGivern, R. F., Huston, P. J., Byrd, D., King, T., Siegle, G. J., & Reilly, J. (1997). Sex differences in visual recognition memory: Support for a sex-related difference in attention in adults and children. *Brain and Cognition*, 34, 323-336. doi:10.1006/brcg.1997.0872
- McGivern, R. F., Mutter, K. L., Anderson, J., Wideman, G., Bodnar, M., & Huston, P. J. (1998). Gender differences in incidental learning and visual recognition memory: Support for a sex difference in unconscious environmental awareness. *Personality and Individual Differences*, 25, 223-232. doi:10.1016/S0919-8869(98)00017-8
- Mintzer, M. Z., & Snodgrass, J. G. (1999). The picture superiority effect: Support for the distinctiveness model. *American Journal of Psychology*, 112, 113-146. doi:10.2307/1423627
- Murdock, B. B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review*, 104, 839-862. doi:10.1037/0033-295X.104.4.839
- Parks, C. M., & Yonelinas, A. P. (2007). Moving beyond pure signal-detection models: Comment on Wixted (2007). *Psychological Review*, 114, 188-202. doi:10.1037/0033-295X.114.1.188
- Prieto, G., & Delgado, A. R. (2003). Análisis de un test mediante el modelo de Rasch. *Psicothema*, 15, 94-100.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Danmarks: Danmarks Paedagogiske Institut.
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11, 251-257. doi:10.1016/j.tics.2007.04.004
- Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, 392, 595-598. doi:10.1038/33396
- Schloerscheidt, A. M., & Rugg, M. D. (2004). The impact of change in stimulus format on the electrophysiological indices of recognition. *Neuropsychologia*, 42, 451-466. doi:10.1016/j.neuropsychologia.2003.08.010
- Stenberg, G. (2006). Conceptual and perceptual factors in the picture superiority effect. *European Journal of Cognitive Psychology*, 18(6), 813-847. doi:10.1080/09541440500412361
- Stenberg, G., Radeborg, K., & Hedman, L. R. (1995). The picture superiority effect in a cross-format recognition task. *Memory and Cognition*, 23, 425-441. doi:10.3758/BF03197244
- Wheeler, M. E., & Buckner, R. L. (2004). Functional-anatomic correlates of remembering and knowing. *NeuroImage*, 21, 1337-1349. doi:10.1016/j.neuroimage.2003.11.001
- Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research*, 1100, 125-135. doi:10.1016/j.brainres.2006.05.019
- Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual process model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 20, 1341-1354. doi:10.1037/0278-7393.20.6.1341
- Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, 25, 747-763. doi:10.3758/BF03211318
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, 46, 441-517. doi:10.1006/jmla.2002.2864
- Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and Cognition*, 5, 418-444. doi:10.1006/ccog.1996.0026
- Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology*, 48, 516-535. doi:10.1037/1196-1961.48.4.516