



Estudos de Psicologia

ISSN: 0103-166X

estudosdepsicologia@puc-  
campinas.edu.br

Pontifícia Universidade Católica de  
Campinas  
Brasil

Cecílio Fernandes, Débora; Prieto, Gerardo; Delgado, Ana Rosa  
Construção de um teste transmodal de memória de reconhecimento  
Estudos de Psicologia, vol. 27, núm. 3, julio-septiembre, 2010, pp. 283-291  
Pontifícia Universidade Católica de Campinas  
Campinas, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=395335744001>

- 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 de um teste transmodal de memória de reconhecimento

## *Construction of a cross-modal recognition memory test*

Débora Cecílio **FERNANDES**<sup>1</sup>

Gerardo **PRIETO**<sup>2</sup>

Ana Rosa **DELGADO**<sup>2</sup>

### Resumo

O objetivo do estudo foi construir um teste de reconhecimento pelo modelo de Rach composto por 64 itens de reconhecimento, 32 itens estudados e 32 distratores. Os itens compreendiam pares de palavras escritas ou de figuras, eram modais ou transmodais e metade pertencia a uma mesma categoria semântica e metade a diferentes categorias. Os pares foram agrupados por essas características e hipotetizou-se que esses grupos seguissem uma determinada ordem de dificuldade e que suas médias de dificuldade se diferenciasssem significativamente. O teste foi aplicado coletivamente em 231 estudantes de uma universidade particular. Os resultados indicaram um bom ajuste dos itens em relação ao *infit* e *outfit* e um bom índice de precisão. A ordem de dificuldade esperada foi obtida empiricamente, mas não foram encontradas diferenças significativas entre as médias de todas as categorias dos itens.

**Unitermos:** Memória. Reconhecimento, aprendizagem. Teoria de resposta ao item.

### Abstract

*The aim of this study was to construct a recognition test, using the Rasch model, composed of 64 items to be recognized, 32 items for study and 32 distracters. The pairs of items comprised written words or figures, they were either modal or cross-modal; half of them belonged to the same semantic category and the other half to a variety of other categories. The pairs were grouped according to these features and it was predicted that they would follow a given order of difficulty and that their mean of difficulty would be significantly different. The test was collectively administered to 231 students from a private university. The outcomes indicated a very good item fit as regards *infit* and *outfit* and also a good index of precision. The expected orders of difficulty were obtained empirically, although the means of some groups of items were not significantly different.*

**Uniterms:** Memory. Recognition learning. Item response theory.

Atualmente a memória é considerada um fenômeno biológico e psicológico que implica a associação de alguns sistemas, circuitos *cerebrais* e

subprocessos mnemônicos que trabalham em conjunto para o funcionamento da memória nos vários aspectos e condições da vida (Emilien, Durlach, Antoniadis, van

▼ ▼ ▼ ▼ ▼

<sup>1</sup> Doutoranda, Universidad de Salamanca, Faculdade Psicologia, Departamento de Psicologia Básica, Psicobiología y Metodología de las Ciencias Del Comportamiento. Av. de la Merced, nº 109-131, Código Postal 37005, Salamanca, Espanã. Correspondência para/Correspondence to: D.C. FERNANDES. E-mail: <decfernandes@hotmail.com>.

<sup>2</sup> Universidad de Salamanca, Faculdade Psicologia, Departamento de Psicologia Básica, Psicobiología y Metodología de las Ciencias del Comportamiento. Salamanca, Espana.

der Linden & Maloteaux, 2004). Um dos sistemas é a memória declarativa, que compreende a memória episódica, que, de modo geral, se refere à memória para eventos passados pessoais que tiveram lugar em um contexto espacial ou temporal específico. Ela é autobiográfica, porque a informação está ligada às experiências pessoais, além de ser composta por eventos que são sequencialmente ordenados no tempo.

A memória episódica pode ser conceituada como a reatuação mental consciente de eventos passados experimentados pessoalmente, ou ainda como a recuperação ou o reconhecimento de um evento único e o contexto espaço-temporal no qual o evento foi experimentado. Além disso, o termo pode ser interpretado como as representações internas que são acessíveis à lembrança. Finalmente, pode ser também o sistema cerebral que dota o sujeito com a capacidade para realizar essas funções (Dudai, 2004).

A informação episódica armazenada está acessível por meio do processo de recuperação, mediante o ato de lembrar ou de reconhecer. O processo de reconhecimento se refere ao julgamento de uma ocorrência prévia ou ao processo cerebral pelo qual esse julgamento é alcançado (Dudai, 2004). Ademais, o reconhecimento pode requerer a associação entre estímulos apresentados ou mesmo entre estímulos e informações contextuais.

Um teste de reconhecimento é a situação na qual o sujeito julga a familiaridade ou quão recente é um estímulo (Dudai, 2004). Nesse contexto, o objetivo deste trabalho foi construir um teste de reconhecimento, utilizando-se o modelo de Rasch (Rasch, 1960), que se refere a um modelo psicométrico cujas propriedades métricas são muito apropriadas para a construção de testes.

Além disso, uma das características do modelo é a medição conjunta de sujeitos e itens, quer dizer, a estimação conjunta tanto dos valores de dificuldade dos itens como os traços dos sujeitos. Geralmente, os modelos com essa característica propõem estimar as diferenças individuais entre sujeitos mediante um conjunto de itens que variam em relação a um atributo latente, nesse caso, a memória de reconhecimento. O modelo de Rasch estima essa variação dos itens mediante o parâmetro de dificuldade.

A variação hipotética da dificuldade dos itens foi obtida a partir de um modelo cognitivo de memória proposto por Hunt e McDaniel (1993), que prevê a superioridade da lembrança ou reconhecimento das figuras em relação às palavras, uma vez que os traços visuais altamente característicos e únicos das figuras que não são compartilhados por outros estímulos são diferencialmente codificados, o que leva a um reconhecimento futuro mais eficaz. Além disso, os autores (Hunt & McDaniel, 1993) também propuseram que a codificação dos elementos estudados e contextuais não é aleatória, mas primeiramente tem lugar a organização das características comuns entre os itens e sua codificação. Posteriormente, a partir dessa organização, são ressaltadas e codificadas as características únicas dos itens (atributos). Os autores consideram que a codificação simultânea tanto das semelhanças como das diferenças dos itens é a condição ótima para um melhor reconhecimento.

Essas considerações levaram à elaboração tanto de itens mais fáceis como de itens mais difíceis. No caso dos itens fáceis, suas características cumpriam as condições para produzir um reconhecimento mais eficaz e, no caso dos itens mais difíceis, seus atributos não proporcionavam as condições para um melhor reconhecimento. Nesse sentido, as propostas dos autores justificaram a inclusão de estímulos tanto no formato de figuras como de palavras escritas, já que prevêm, assim como vários estudos, que as figuras são mais facilmente reconhecidas que as palavras escritas (Mintzer & Snodgrass, 1999; Stenberg, Radeborg & Hedman, 1995).

O estudo das figuras é necessário devido ao papel delas em algumas estratégias de aprendizagem e de retenção de informação, que são de suma importância para a vida cotidiana, mas, principalmente, porque oferecem formas alternativas para confrontar algumas condições que afetam gravemente a memória. É o caso, por exemplo, do envelhecimento normal ou com deterioro cognitivo relacionado com a idade, assim como de pacientes que apresentam quadros de amnésia. Nesse sentido, o estudo realizado por Campos, Gómez-Juncal e Pérez-Fabello (2008) mostrou que o uso de imagens como estratégia de aprendizagem melhorou o desempenho a longo prazo (após um dia e sete dias) dos participantes em testes de lembrança livre e de

reconhecimento. Além disso, os dados também mostraram que os participantes que possuíam experiência com o uso de imagens mentais como estratégia apresentaram um desempenho superior em comparação com os sujeitos que não tinham esse tipo de experiência. Ademais, essa superioridade foi observada em períodos de um a sete dias após o contato com os itens estudados (Campos et al., 2008).

Outra característica dos itens do teste reconhecimento consiste em que eles são compostos por pares de estímulos, de modo que os itens requerem um processo de reconhecimento associativo, que se trata de uma tarefa mais complexa, porque exige o reconhecimento da associação dos pares de estímulos apresentados e não o reconhecimento de cada item individualmente (Figuras 1 e 2). Os pares de estímulos estudados eram, portanto, apresentados em formato de figura ou palavra escrita e os seus membros podiam pertencer a uma mesma categoria semântica ou a categorias distintas. A semelhança entre pares estudados facilita sua futura recuperação de modo que pares de estímulos que pertencem a uma mesma categoria semântica são mais facilmente recuperados que pares que não guardam nenhuma relação semântica entre si (Hintzman, 1988).

Uma das variáveis manipuladas para variar o nível de dificuldade dos itens foi a modalidade/transmodalidade. Por um lado, a condição de modalidade implica estudar e reconhecer os itens em um mesmo formato, por exemplo, estudar figuras e reconhecer figuras. Por outro lado, a transmodalidade é a situação na qual um sujeito estuda o estímulo em um determinado formato para posteriormente reconhecê-lo em outro formato, como estudar pares de figuras e, na situação de teste, reconhecê-las como palavras.

Nesse contexto, as pesquisas que manipulam as diferentes modalidades sensoriais em tarefas de reconhecimento indicam que a transmodalidade as dificulta (Mintzer & Snodgrass, 1999; Schloerscheidt & Rugg, 2004; Stenberg et al., 1995). Existem evidências que sustentam que quando o formato de codificação dos estímulos são figuras, mas o formato de reconhecimento são palavras correspondentes às figuras (FP), dificulta-se o desempenho da tarefa de reconhecimento. Contudo, essa dificuldade é ainda maior quando se trata de codificar palavras e reconhecê-las como figura (PF),

sendo esta a condição de transmodalidade que mais dificulta o reconhecimento.

Devido ao efeito da superioridade das figuras, considera-se que a codificação é favorecida quando o estímulo a ser codificado é uma figura, mas se essa figura sofre o efeito da transmodalidade (deve ser reconhecida como palavra, FP), a sua dificuldade se vê um pouco aumentada. Por isso, a condição FP, devido à transmodalidade, deveria apresentar igual dificuldade que a condição modal de codificação de uma palavra e futuro reconhecimento como palavra (PP), já que essa condição, apesar de não se favorecer de uma codificação diferenciada, também não se prejudica pela transmodalidade (Mintzer & Snodgrass, 1999; Schloerscheidt & Rugg, 2004; Stenberg et al., 1995).

Então, a variação dos níveis de dificuldade dos itens se baseou nessas três variáveis: formato de codificação (figuras ou palavras), presença/ausência de categoria semântica e modalidade/transmodalidade. Do mesmo modo, outras variáveis foram mantidas constantes com o intuito de evitar suas influências no reconhecimento. Por isso, foram selecionadas somente palavras bissílabas e figuras cujas palavras correspondentes eram bissílabas. Tanto as palavras como as figuras escolhidas eram de uso frequente ou muito frequente, já que, ao contrário da livre lembrança, o reconhecimento é facilitado por palavras de pouca frequência de uso (Anderson & Bower, 1972). Ademais, foi controlada a concreção das palavras, optando-se por palavras concretas em vez de palavras abstratas, pois o uso de palavras abstratas não permitiria a utilização de figuras correspondentes. Por tanto, as variáveis frequência de uso, tamanho e concreção das palavras foram controladas para exercer o mínimo de influência possível no reconhecimento, para tentar garantir que o desempenho fosse afetado unicamente pelas variáveis modalidade/transmodalidade, associação/não associação semântica e formato de codificação dos estímulos.

Além desses itens de estudo, foram construídos, também, itens distratores cujo objetivo era aumentar a dificuldade geral do teste. A sua configuração se baseou em estudos de falso reconhecimento, apesar de não se terem dado as condições ideais para produzir esse efeito. Uma das maneiras de gerar falsos reconhecimentos é por meio da apresentação de uma lista de estudo alta-

mente estruturada com pares de itens muito semelhantes durante a fase de estudo e, durante o teste de reconhecimento, trocar alguns membros dos pares por outros estímulos diferentes ao membro original, mas caracterizados pelas mesmas semelhanças. Contudo, a lista desse teste de reconhecimento se caracterizava por ser mista, ou seja, não era altamente estruturada, mas composta por alguns itens de uma mesma categoria semântica e outros itens sem nenhuma semelhança aparente. Portanto, os membros de alguns pares foram trocados por outros relacionados semanticamente, mas não se levantaram hipóteses sobre eles devido ao tipo de lista de estudo desse teste. Unicamente se esperou que os distratores aumentassem a dificuldade geral do teste pelo aumento da quantidade de itens no teste de reconhecimento.

Não obstante, para os itens estudados foram levantadas hipóteses acerca de seus níveis de dificuldade, considerando as possibilidades de interação entre as três variáveis manipuladas. Esperou-se que a média de dificuldade de cada categoria de itens se diferenciasse significativamente das demais e que seguissem uma determinada ordem de dificuldade de acordo com a interação das variáveis dos itens de cada categoria. Nesse contexto, as categorias e a ordem crescente de dificuldade dos itens foram: categoria A (pares de figura-figura associados), categoria B (pares de figura-figura não associados), categoria C (pares de figura-palavra associados e palavra-palavra associados), categoria D (pares de figura-palavra não associados e palavra-palavra não associados), categoria E (pares de palavra-figura associados) e categoria F (pares de palavra-figura não associados). Além disso, esperou-se que o conjunto de itens elaborados gerasse diferenças individuais entre os sujeitos, ou seja, que ao entrar em contato com eles os sujeitos manifestariam suas habilidades de reconhecimento e que, por isso, os itens seriam úteis para medir o constructo da memória de reconhecimento.

## Método

### Modelo de Rasch

O modelo de Rasch (Rasch, 1960) é um modelo probabilístico de um parâmetro que parte da premissa fundamental de que o comportamento de um indivíduo

ao enfrentar um item pode ser explicado em função da sua habilidade latente. Apesar de que essa habilidade não é diretamente observável, pode ser medida pelas situações de testes que propiciam sua manifestação (Bond & Fox, 2001).

O modelo de Rasch calcula a probabilidade de acerto de cada sujeito para cada item e, a partir desses valores, estima conjuntamente os parâmetros de dificuldade dos itens e de habilidade dos sujeitos. A fórmula da probabilidade de acerto é:

$$P(\chi=1, \beta, \delta) = \frac{e^{(\beta, \delta)}}{1 + e^{(\beta, \delta)}}$$

sendo  $\beta$  corresponde à dificuldade dos itens e  $\delta$  ao parâmetro de habilidade dos sujeitos. É possível observar que a probabilidade de acerto de um item depende somente dessas variáveis e a diferença entre o nível do traço de um sujeito e a dificuldade do item tem um significado direto em relação a essa probabilidade (Embretson & Reise, 2000).

Do mesmo modo, esses dois parâmetros estão completamente separados, o que possibilita que a estimação de um deles não dependa do valor do outro, mas sim de seus atributos. Por isso, a estimação da dificuldade de um item depende exclusivamente dos atributos do item e não das características dos sujeitos (Wright & Mok, 2004), o que implica que os valores estimados sejam invariáveis.

O modelo de Rasch estima simultaneamente ambos os parâmetros, por meio de métodos de máxima verossimilhança. A medição conjunta independe da amostra de sujeitos que responderam aos itens, e, por isso, os valores dos parâmetros podem ser extrapolados e podem gerar expectativas de respostas futuras.

As inferências, generalizações e propriedades fornecidas pelo modelo são possíveis apenas quando os dados obtidos se ajustam às expectativas do modelo (Bond & Fox, 2001). Nesse sentido, o modelo de Rasch utiliza o *infit* e o *outfit* como estatísticos de ajuste que diferem entre si com respeito às técnicas de medição do ajuste. No caso do *infit*, ele considera mais o desempenho dos sujeitos cuja habilidade está próxima à dificuldade do item, enquanto o *outfit* é mais sensível às influências das pontuações anômalas e às respostas de sujeitos aos itens distantes de seu nível de habilidade (Wright & Mok, 2004).

## Participantes

Participaram da pesquisa 231 sujeitos, a maioria dos quais era estudante de psicologia ( $n=168$ ) de uma universidade particular do interior paulista. O restante do grupo abarcava professores, estudantes de matemática, pedagogia, arquitetura, economia e de mestrado e de doutorado. A idade média dos sujeitos foi de 25 anos Desvio-padrão (DP)=7 e a porcentagem de mulheres foi de 81%. O estudo faz parte do projeto aprovado pelo Comitê de Ética da Universidade São Francisco, protocolo número 209/02, em 2002; todos os participantes leram e assinaram o termo de consentimento livre e esclarecido antes de sua inclusão na amostra.

## Instrumentos

O teste construído consiste em duas partes fundamentais. A primeira se refere à apresentação em *Power Point* de todas as instruções do teste e dos itens da lista de estudo. As instruções dadas ensinam como responder os itens do teste, incluindo exemplos dos tipos de estímulos presentes (pares de palavras ou de figuras), exemplos de estudo dos estímulos e a subsequente resolução dos itens, reproduzindo exatamente a situação do teste. Além disso, nessa parte do teste, a lista de estudo, composta por 32 pares de estímulos, é apresentada duas vezes seguidas, mas na segunda exibição, a apresentação tem início a partir do item 17 e segue a mesma ordem até o último item, que é correspondente ao 16 da primeira apresentação.

Já a segunda parte do teste corresponde ao caderno de respostas. Na primeira página os sujeitos devem preencher os seus dados pessoais; na página seguinte está o exemplo de resolução do teste, constituído por seis pares de estímulos. A terceira página contém as respostas corretas dos exemplos marcadas com um X.

Nas páginas posteriores se encontra o teste de reconhecimento, constituído por 64 itens; cada um deles contém sete opções de respostas formadas por palavras, figuras e a opção "o par não está presente". Os 32 pares estudados podem estar presentes no teste segundo as condições de modalidade e transmodalidade e os demais 32 itens são distratores, contêm somente um dos membros dos pares estudados, podendo esse

membro estar sozinho ou acompanhado de outro estímulo da mesma categoria semântica. Também contém sete opções de respostas. Segue abaixo um exemplo de um dos possíveis itens (Figuras 1 e 2).

Todas as palavras incluídas no teste são concretas, dissílabas e de uso frequente ou muito frequente. No caso das figuras usadas, suas palavras correspondentes também possuem essas mesmas características.

## Procedimentos

A aplicação foi feita coletivamente, nas salas de aula e durou 20 minutos. As instruções e a lista de estudo foram apresentadas com o *Data Show*. Em primeiro lugar, os aplicadores expuseram a pesquisa e foi explicado o caráter voluntário da participação; foi entregue e lido pelo aplicador o termo de consentimento livre e esclarecido, e os alunos que se dispuseram a responder ao instrumento assinaram as duas vias do termo. Os alunos que participariam da pesquisa receberam o material e preencheram a página com os dados pessoais. A seguir, foi iniciada a apresentação do *Power Point* com as instruções, exemplos e a lista de estudo.

A análise dos dados foi dividida em duas fases, nas quais foram realizadas as mesmas análises. Na primeira fase foram considerados os 64 pares de itens, enquanto na segunda fase foram excluídos os 32 itens distratores. Em ambas as fases foram utilizados para as análises os programas *Winsteps* (Linacre, 2002) e *SPSS*.

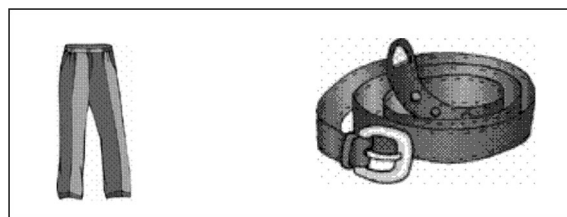


Figura 1. Par de estímulos da lista de estudo.



Figura 2. Item do teste de reconhecimento.

## Resultados

### Fase 1

Nessa fase, o índice de precisão foi bastante adequado tanto para os sujeitos (0,92) como para os itens (0,97); com relação ao valor do erro de medida dos sujeitos, sua média foi de 0,36 (DP=0,9) e dos itens foi de 0,18 (DP=0,03). Além disso, a média da habilidade dos sujeitos foi de 1,15.

Para os itens, a média do *infit* foi de 0,98 (DP=0,24) e a média do *outfit* foi de 1,06 (DP=0,39). Com respeito aos sujeitos, a média do *infit* obtida foi de 0,99 (DP=0,22), enquanto a do *outfit* foi de 1,06 (DP=0,53). Uma análise mais detalhada dos itens possibilitou verificar quais itens são úteis para a medição do construto, pois estão dentro do intervalo de 0,5 e 1,5 (Linacre, 2002). Por isso, em relação ao *infit* deve ser excluído o item A26, e com respeito ao *outfit*, devem ser eliminados seis itens: A26, C3, B15, A12, B25, C56.

Em relação à adequação dos sujeitos, três deles apresentaram valores de *infit* desajustados. Ao lado disso, 31 pessoas obtiveram valores de *outfit* considerados por Linacre (2002) como padrões de resposta com excessivo desajuste.

Em último lugar, os valores mínimos e máximos de dificuldade dos itens indicaram que o intervalo de distribuição ao longo da escala *logit* variou de -2,56 a +3,56, tendo a maioria se agrupado entre os valores de -1 a +1. Na Figura 3 é possível visualizar a distribuição dos itens e dos sujeitos ao longo da escala *logit*.

Com relação à análise das seis categorias de dificuldade, a Tabela 1 fornece seus valores descritivos. A categoria A foi a mais fácil de todas, com valor médio de dificuldade de -1,78, seguida da categoria B; as categorias C e D foram consideradas de dificuldade medianas e as E e F, as mais difíceis.

A Análise de Variância (ANOVA) dos valores de dificuldade das categorias indicou [ $F(5,231)=22,95$ ,  $p=0,000$ ], o que sugere que as diferenças encontradas entre as médias das categorias não podem ser atribuídas ao acaso, isto é, são estatisticamente significativas. A prova *a posteriori* de Bonferroni revelou que não foram encontradas diferenças significativas entre todas as seis categorias. A categoria F foi a única significativamente diferente de todas as categorias.

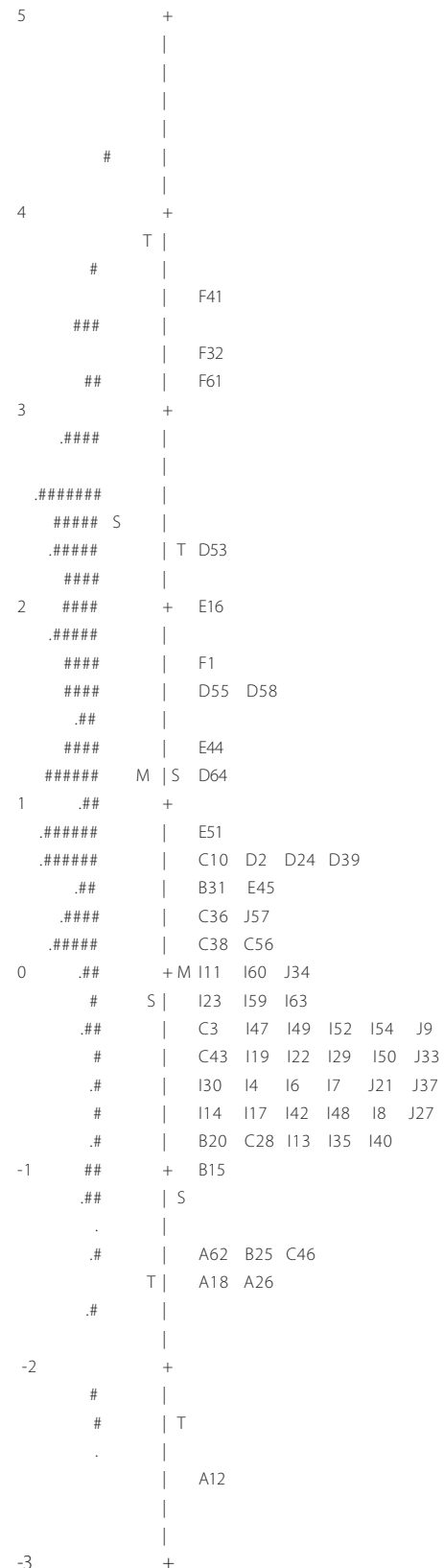


Figura 3. Mapa dos sujeitos e itens.

Nota: cada '#' são 2 sujeitos.



**Tabela 1.** Média e desvio-padrão das categorias de dificuldade da Fase 1.

Categorias	Média	Desvio-Padrão
A	-1,78	0,53
B	-0,73	0,87
C	-0,15	0,74
D	1,12	0,67
E	1,21	0,60
F	2,89	0,82
<b>Total</b>	0,44	1,51

Finalmente, a correlação ordinal de Spearman mostrou uma correlação perfeita entre a ordem de dificuldade esperada das categorias dos itens e a ordem obtida empiricamente. Essa análise indica, portanto, que a ordem crescente de dificuldade das categorias foi A, B, C, D, E e F.

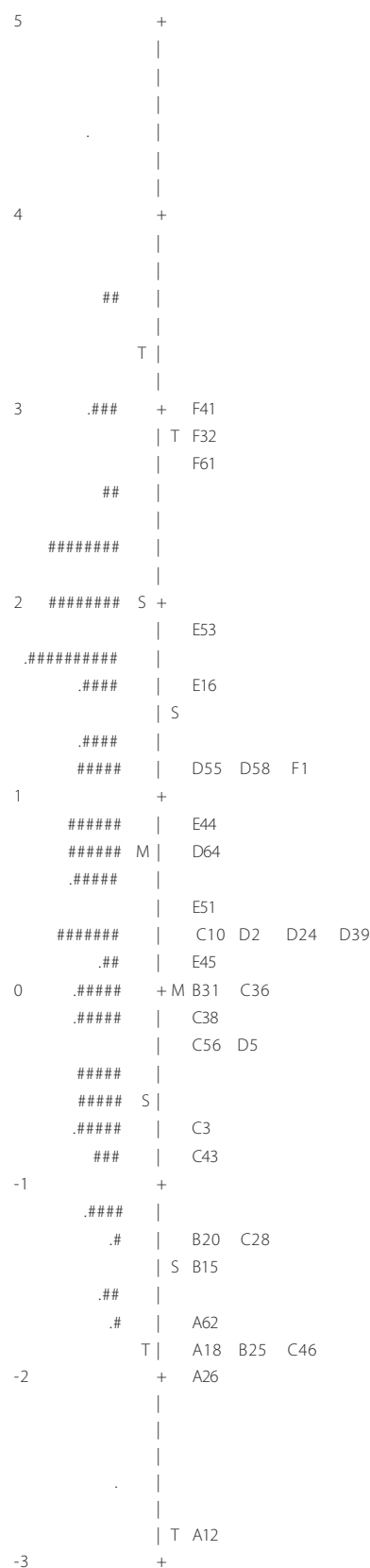
## Fase 2

Após a eliminação dos itens distratores, a média da habilidade dos sujeitos foi de 0,72, menor que a anterior, o que indica que os distratores facilitaram o desempenho dos sujeitos no teste de reconhecimento. Outra variação interessante é que, por um lado, ao diminuir pela metade o número de itens para a estimação das habilidades dos sujeitos, a precisão dos sujeitos também foi reduzida, passando de 0,92 a 0,84; por outro lado, a precisão dos itens aumentou para 0,98.

Com relação ao ajuste dos itens e dos sujeitos, houve pouca mudança nos seus valores: a média do *infit* dos sujeitos foi de 1,00 (DP=0,25), ao passo que a média obtida do *outfit* foi 1,09 (DP=0,83). No que se refere aos itens, a média do *infit* foi 1,00 (DP=0,11) e a do *outfit* foi de 1,09 (DP=0,35).

De acordo com o critério de Linacre (2002), nenhum item deve ser excluído ao considerar os valores de *infit*, mas quatro itens apresentaram desajuste em relação ao *outfit*: A26, C3, B15 e A18. Já os valores de ajuste dos sujeitos indicaram que seis sujeitos não se ajustaram devido aos valores de *infit* e 26 devido ao *outfit*.

Apesar da exclusão dos 32 distratores, os itens estudados se distribuem consideravelmente ao longo da escala *logit*, e esse conjunto de itens é adequado para essa amostra de sujeitos (Figura 4).



**Figura 4.** Mapa dos sujeitos e itens.

Nota: cada '#' são 2 sujeitos



Em relação à análise da dificuldade das categorias, cujos valores podem ser observados na Tabela 2, a categoria A, considerada a mais fácil, teve uma média de -2,11, seguida da categoria B. As categorias C e D obtiveram médias de dificuldade de -0,60 e 0,64, respectivamente, o que as coloca como de mediana dificuldade. Finalmente, as duas categorias difíceis foram a categoria E, com uma média de 0,73, e F que obteve uma média de 2,43 (Tabela 2).

A análise de variância indicou [ $F(5,231)=23,00$ ,  $p=0,000$ ], o que significa que as diferenças encontradas entre as médias não podem ser atribuídas ao acaso. A prova *a posteriori* de Bonferroni revelou que também não houve diferença significativa entre as médias de todas as categorias, isto é, que não foram encontrados seis níveis de dificuldade bem definidos. Aqui, as diferenças significativas entre as médias das categorias de itens foram as mesmas que na fase anterior. Finalmente, a correlação ordinal de Spearman uma vez mais indicou a existência de uma correlação perfeita entre a ordem esperada de dificuldade das categorias e sua ordem obtida empiricamente.

## Discussão

De modo geral, os valores dos indicadores de precisão e adequação dos itens foram bastante bons para ambas as fases de análise, o que sugere que os itens se adequaram às expectativas do modelo de Rasch e que por isso é possível aceitar as inferências, generalizações e propriedades do modelo acerca dos dados obtidos. A maioria dos itens favorece a medição da memória de reconhecimento, segundo os critérios de Linacre (2002) e de acordo com os indicadores de pre-

cisão, que foram bastante adequados: a replicabilidade dos parâmetros estimados está assegurada.

O mapa dos itens e sujeitos da Fase 1 mostra claramente um excesso de itens dentro do intervalo de -1 a 1. Apesar dessa aglomeração, existe uma distribuição de itens considerável ao longo da escala *logit*, de modo que essa variabilidade no nível de dificuldade dos itens mede com certa precisão os diferentes níveis de habilidade dos sujeitos. Essa capacidade para produzir diferenças individuais é uma condição necessária para considerar que a escala tem boas qualidades psicométricas e é útil para a medição do constructo.

A Fase 2, cujas análises excluíam os 32 distratores, mostrou que o aumento esperado da dificuldade geral do teste proporcionado por esses itens não foi atingido, já que o desempenho dos sujeitos aumentou em comparação com a fase anterior. Além disso, a sua eliminação provocou uma pequena melhora dos índices de ajuste e precisão dos itens estudados, o que sugere que os 32 itens estudados medem de forma mais precisa o constructo memória de reconhecimento, em comparação com análise dos 64 itens.

A correlação ordinal de Spearman mostrou que a ordem esperada de dificuldade das categorias foi obtida empiricamente, de modo que a interação entre as variáveis formato de codificação, modalidade/transmodalidade e presença/ausência de categoria semântica teve como resultado a variação da dificuldade dos itens de acordo com as hipóteses levantadas. Nesse sentido, os pares de figuras (modais) foram mais facilmente reconhecidos que os de palavras (modais), assim como os pares associados semanticamente foram mais fáceis que os pares sem associação semântica e, finalmente, os pares modais foram mais fáceis que os pares transmodais. Apesar disso, as diferenças entre as categorias de itens não formaram seis grupos claramente definidos, já que não houve diferença significativa entre as médias de todas elas.

O reconhecimento é uma habilidade mais fácil que outros tipos de memória (Craik, 1986; Dudai, 2004) e essa facilidade muitas vezes pode produzir o efeito teto, que consiste em que a maioria dos sujeitos alcance a pontuação máxima do teste. Nesse sentido, o presente teste evitou esse efeito e possibilitou observar as diferenças individuais ao responder os itens e contemplar como cada categoria de itens se localizou em diferentes

**Tabela 2.** Média e desvio-padrão das categorias de dificuldade da Fase 2.

Categorias	Média	Desvio-Padrão
A	-2,11	0,49
B	-1,14	0,88
C	-0,60	0,69
D	0,64	0,67
E	0,73	0,60
F	0,84	0,84
<b>Total</b>	0,00	1,47

pontos da escala *logit* de acordo com suas características.

Um ponto interessante dos itens desse teste é a presença da transmodalidade, que aumentou a dificuldade dos itens em comparação com os itens modais, corroborando os estudos de Mintzer e Snodgrass (1999), Schloerscheidt e Rugg (2004) e Stenberg et al. (1995). Os três experimentos apresentaram o mesmo padrão de respostas, caracterizado por uma maior quantidade de erros, um tempo de reação mais lento e uma menor precisão no desempenho de itens transmodais ao ser comparado com o desempenho em itens modais.

## Considerações Finais

Cabe comentar que esse teste de reconhecimento ainda necessita mais estudos, uma revisão ou eliminação dos distratores e dos itens que não se ajustaram ao modelo de Rasch. Ademais, seria interessante incluir itens que avaliassem melhor os sujeitos que mostram uma habilidade superior, e, em estudos futuros, averiguar se os diferentes padrões de respostas de populações clínicas se veem refletidas nesse teste, assim como comparar os resultados das aplicações desse instrumento com outros instrumentos de memória com o fim de observar as possíveis relações entre eles. Finalmente, após as modificações necessárias, esse teste deverá ser informatizado.

## Referências

- Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, 79 (2), 97-123.
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: fundamental measurement in the human sciences*. New Jersey: Lawrence Erlbaum.
- Campos, A., Gómez-Juncal, R., & Pérez-Fabello, M. J. (2008). Experiencia en la mnemotecnica y aprendizaje incidental con imágenes normales y raras. *Estudios de Psicología*, 25 (3), 321-331.
- Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp.409-422). Amsterdam: Elsevier.
- Dudai, Y. (2004). *Memory from A to Z: keywords, concepts and beyond*. New York: Oxford University Press.
- Embretson, S. E., & Reise, S. P. (2000). The trait level measurement scale: meaning, interpretations, and measurement-scale properties. In S.E. Embretson & S.P. Reise. *Item response theory for psychologists* (pp.125-157). New Jersey: Lawrence Erlbaum.
- Emilien, G., Durlach, C., Antoniadis, E., van der Linden, M., & Maloteaux, J. M. (2004). *Memory: neuropsychological, imaging, and psychopharmacological perspectives*. New York: Psychology Press.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95 (5), 528-551.
- Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, 32 (4), 421-445.
- Linacre J. M. (2002). What do Infit and Outfit, Mean-Squared and Standardized mean? *Rasch Measurement Transactions*, 16 (2), 878. Recuperado em julho 10, 2006, disponível em: <http://209.238.26.90/rmt/rmt82a.htm>
- Mintzer, M. Z., & Snodgrass, J. G. (1999). The picture superiority effect: support for the distinctiveness model. *American Journal of Psychology*, 112 (1), 113-146.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danmarks Paedagogiske Institut.
- Schloerscheidt, A. M., & Rugg, M. D. (2004). The impact of change in stimulus format on the electrophysiological indices of recognition. *Neuropsychologia*, 42 (4), 451-466.
- Stenberg, G., Radeborg, K., & Hedman, L. R. (1995). The picture superiority effect in a cross-format recognition task. *Memory and Cognition*, 23 (4), 425-441.
- Wright, B. D., & Mok, M. M. C. (2004). An overview of the family of Rasch measurement models. In: E. V. Jr. Smith & R. M. Smith. *Introduction to Rasch measurement: Theory, models and applications* (pp.1-24). Maple Grove: JAM Press.

Recebido em: 4/10/2007

Versão final reapresentada em: 23/3/2010

Aprovado em: 20/5/2010