



Psicologia: Reflexão e Crítica

ISSN: 0102-7972

prcrev@ufrgs.br

Universidade Federal do Rio Grande do Sul  
Brasil

Andriola Bandeira, Wagner

Funcionamento Diferencial dos Itens (DIF): Estudo com Analogias para Medir o Raciocínio Verbal

Psicologia: Reflexão e Crítica, vol. 13, núm. 3, 2000, pp. 475-483

Universidade Federal do Rio Grande do Sul

Porto Alegre, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=18813315>

- ▶ 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

## Funcionamento Diferencial dos Itens (DIF) Estudo com Analogias para Medir o Raciocínio V

*Wagner Bandeira Andriola*  
Universidade Federal do Ceará

### Resumo

Este estudo objetivou determinar o funcionamento diferencial de 30 analogias destinadas à avaliação considerando a variável sexo. Utilizou-se uma amostra de 730 alunos do Ensino Médio, com idade média de 17,74 anos ( $sd = 3,12$  anos). A maioria procedia de escolas públicas (58,5%) e era do sexo feminino (53,2%). Os grupos investigação foram compostos por homens ( $n=342$ ) e mulheres ( $n=388$ ). Os parâmetros métricos dos itens foram determinados pelo modelo TRI de dois parâmetros logísticos. Para a verificação do DIF foram comparados os parâmetros de DIF. Os resultados indicaram a presença de cinco itens com DIF.

*Palavras-chave:* Funcionamento diferencial dos itens (DIF); teoria da resposta ao item (TRI); raciocínio verbal; psicológica.

**Differential Items Functioning (DIF): Study with Analogies for Measurement the Verbal Reasoning**

### Abstract

This research aimed the determination of the differential item functioning (DIF) in 30 analogies used for assessment in students, taking into account the sex variable. A sample of 730 high school students, with a mean age of 17,74 years ( $sd = 3,12$  years) was used. The majority was composed by students from public schools (53,3%). The groups which participated in the study of DIF were composed by men ( $n= 342$ ) and women ( $n=388$ ). The metric parameters of the items were determined according to the TRI model of two logistics parameters of the DIF the method of comparison of the metric parameters of the items was used. The results indicated five items with DIF.

*Keywords:* Differential items functioning (DIF); item response theory (IRT); verbal reasoning; psychological.

No âmbito da Teoria Clássica dos Testes (TCT) o termo *viés* é utilizado para rotular os itens que possuem parâmetros de dificuldade ou discriminação diferentes nos distintos grupos investigados. De acordo com Camilli e Shepard (1994), o viés é uma fonte de invalidez ou de erro sistemático que influencia no modo como um teste ou item mede aos membros de um grupo particular; é sistemático porque produz distorções nos resultados de

– O item tem viés se existe diferença entre os índices de dificuldade entre populações distintas;  
– O item tem viés se existe diferença entre os rendimentos entre os grupos de teste ao escore total no teste e na escala;

– O item tem viés quando

sociodemográfica dos sujeitos. Neste âmbito, sempre que se identifiquem grupos de sujeitos para os quais haja suspeita de diferenças nas pontuações obtidas no teste, se deve aplicar algum procedimento para a análise de potenciais viéses (Martínez Arias, 1997).

Hambleton (1989a), ao finalizar sua exposição sobre as várias definições de viés, constata que todas elas padecem de um problema comum: não consideram a necessidade de controlar a própria capacidade dos sujeitos na variável latente medida pelo item ou teste. Reside aqui a principal diferença entre os conceitos de *viés* e de *funcionamento diferencial* dos itens (*DIF*). O procedimento DIF trata de controlar a magnitude da variável latente (geralmente expressa pela letra grega  $\theta$ ) no item avaliado, isto é, os grupos são comparados com respeito às suas pontuações no item ou teste, considerando-se que suas magnitudes na variável latente (capacidades, habilidades ou aptidões) têm idêntico valor (Hambleton, 1994).

Alguns autores, entre os quais Camilli e Shepard (1994), insistem que os índices estatísticos utilizados na análise do DIF por si mesmos não proporcionam prova de viés, preferindo denominá-los de *índices de discrepância do item* ou de *funcionamento diferencial do item*. Segundo eles, este último termo engloba os diferentes procedimentos estatísticos para a detecção de um possível funcionamento diferencial, todavia, insistem em que o DIF não é sinônimo de viés, embora alguns autores creiam que sim. Desse modo, os termos *funcionamento diferencial do item* (DIF) e *viés* não deveriam empregar-se como sinônimos.

É que os métodos estatísticos de DIF se utilizarão para identificar itens que exibem um funcionamento diferencial nos distintos grupos e, posteriormente, depois de uma análise lógica ou experimental no contexto da validade de construto dos itens, se determinará quais têm viéses para que, assim, possam ser eliminados do teste ou banco de itens. Em outras palavras, os métodos de

### **Definição do Termo “Funcionamento do Item” (DIF)**

Atualmente, quase não se utiliza o termo *viés*, que foi preterido pelo de *funcionamento do item*. Uma das razões que explicam tal troca é a existência de técnicas desenvolvidas para a detecção do DIF que proliferado muito nos últimos anos, Estados Unidos de América.

Outra razão deve-se à potência e facilidade para o estudo do DIF, já que foram desenvolvidas técnicas que permitem detectar se um item funciona diferentemente para grupos de distintas características sociodemográficas, cujos sujeitos compõem a mesma magnitude na variável medida (Penfield, 1997; Jiang & Stout, 1998; Kim, Oshima, Raju & Flowers, 1997; Scheuerlein, 1997; Wainer & Lukhele, 1997; Willian, 1997). No entanto, é importante lembrar que, tudo, as técnicas para a detecção do DIF só conseguem detectar se um item funciona mais ou menos que outros, quer dizer, não possuem a capacidade de fornecer informações acerca de sua natureza ou causa (Baker, 1993; Muñiz, 1997).

Cohen, Kim e Baker (1993) distinguem entre o DIF quanto aos objetivos pretendidos e as investigações para a detecção do DIF. No primeiro caso, os autores sugerem que as investigações que utilizam algum método estatístico para a detecção do DIF, a identificação do DIF. Neste caso, os estatísticos são utilizados para detectar somente a detecção do DIF, quer dizer, a existência de uma possível diferença entre as CCI's dos itens avaliados para os grupos comparados. No segundo caso, as investigações para a busca e identificação do DIF. Neste contexto, o objetivo do investimento é a busca de técnicas para a detecção do DIF, para que se possa determinar quais são as causas (psicológicas, educacionais, sociais, atitudinais, etc.) que ocasionam o funcionamento diferencial de um item (Downing & Hembree, 1992).

No âmbito da TRI um item não tem viés se a curva característica do item (CCI) é idêntica para todos os grupos de sujeitos.

–  $T_{jGR}$  é a pontuação verdadeira do sujeito  $j$  que pertence ao grupo de referência e que tem uma certa magnitude na variável latente  $\theta$ ;

–  $T_{jGF}$  é a pontuação verdadeira do sujeito  $j$  que pertence ao grupo focal e que tem uma certa magnitude na variável latente  $\theta$ .

De acordo com Mazor, Hambleton e Claeuser (1998), o uso do número de respostas corretas para a determinação do DIF só é aceitável no caso do teste ser unidimensional e, ademais, quando as respostas são dicotômicas. Como modo de visualizar o DIF em um hipotético item se apresenta a Figura 1.

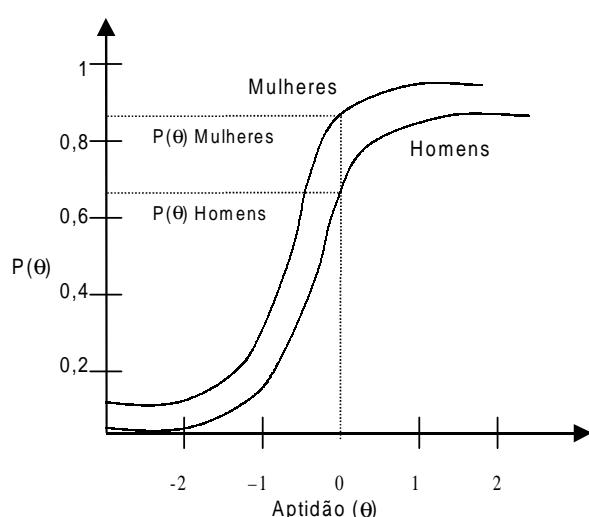


Figura 1. Representação gráfica das CCI's de um item com DIF

Podemos observar que para uma mesma magnitude de  $\theta$  o valor de  $P(\theta)$  é sempre superior para as mulheres, ou seja, em níveis iguais de competência na variável medida  $\theta$  não existe a mesma probabilidade de superar o item. Neste caso, o item tem viés contra os homens (GR) pois

Ainda utilizando o item da Figura 1, o que ocorre com um item sem DIF, tanto, a Figura 2.

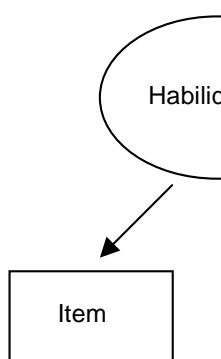


Figura 2. Relação entre habilidade e item sem DIF

O círculo indica a habilidade. Se a habilidade tem uma relação causal com a variável latente, os dois estão associados. Em um exemplo de exemplificar, poderíamos dizer que se uma pessoa tem elevada habilidade na variável latente, ela tem uma relação causal com o grupo com maior capacidade. Se a variável latente tem mais respostas corretas, o rendimento no item depende da magnitude da variável latente. Ou seja, se trata de um item sem DIF.

Agora, observemos o que ocorre quando é representado na Figura 3.

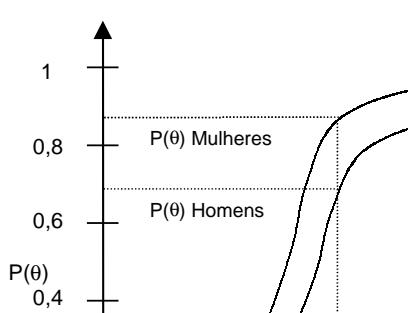
pode favorecer o rendimento de um grupo sobre outro devido a características particulares como sexo, raça, *background* educativo, origem social, etc. Deve ser enfatizado que, neste caso, se supõe que a magnitude da variável latente está sendo controlada, ou seja, os sujeitos são comparados com respeito a seu rendimento tendo em conta que possuem a mesma magnitude no construto. Este segundo exemplo caracteriza o caso no qual o rendimento no item não depende somente da magnitude da variável latente que os indivíduos tenham, senão de características do grupo, isto é, se trata de um item com DIF. Em nosso exemplo, a característica do grupo que influencia o rendimento diferenciado no item é o fato do sujeito ser homem ou mulher. Em síntese, se trata de uma característica de natureza demográfica que influencia o rendimento dos sujeitos com a mesma capacidade (Andriola, no prelo a).

É necessário reconhecer que o DIF ocasiona sérias implicações ao processo de avaliação, já que pode privilegiar um determinado grupo em detrimento de outro (Douglas, Roussos & Stout, 1996), conforme observamos no exemplo do rendimento dos homens e das mulheres. Muñiz (1997) adverte que o problema pode ter repercussões sociais mais graves se é a cultura dominante que elabora os itens para avaliar os sujeitos oriundos de culturas minoritárias. Por exemplo, suponhamos que são construídos itens para avaliar a capacidade de raciocínio verbal em alunos de escolas públicas e privadas. Ocorre que os alunos destes tipos de escolas são, geralmente, oriundos de classes sociais muito distintas, com diferentes *backgrounds* culturais, sociais, econômicos, etc. Todos estes aspectos podem implicar em que um tipo de aluno tenha o vocabulário mais rico que o outro. Dado que o raciocínio verbal é medido através de itens que empregam palavras, muito provavelmente, aquele tipo de aluno que conheça melhor o vocabulário utilizado nos itens terá uma clara vantagem na resolução destes mesmos itens.

Como nos fala Muñiz (1997), é interminável e pode dizer-se que não existem itens isentos completamente de viéses. Na verdade, trata de detectar a quantidade de viéses que existem em um determinado teste ou item. Finalmente, é importante enfatizar que neste contexto, a importância dos itens que objetivam verificar a existência de DIF é justificada. Cabe ao avaliador verificar se existem itens com DIF para que possa elaborar explicações e, assim, evitar sua utilização em desvantagem (Hambleton, 1989b; Lord et al., 1994).

#### Tipos de DIF no Âmbito da TRI

Enfatizamos que no contexto da TRI, a detecção do DIF consiste na comparação entre os resultados obtidos em um item entre os grupos investigados, considerando os grupos de referência e focal. Os distintos métodos para a detecção de DIF foram desenvolvidos baseados em dois tipos principais. O primeiro tipo é denominado *DIF uniforme*, que se observa quando as CCI's do item são paralelas em relação ao nível de rendimento, respeito aos grupos de referência e focal, mas não se cruzam. Em outras palavras, existe uma vantagem relativa para um dos grupos, cujo valor é constante ao longo de todo o intervalo de rendimento. Este caso ocorre quando o item não discrimina (discriminação) tem o mesmo valor na referência e no focal. O segundo tipo é, quando as CCI's são paralelas. Este caso é representado na Figura 4.



homens (grupo de referência), indicando que o item é mais fácil para o grupo focal. Tal diferença no parâmetro  $b$  (dificuldade) supõe que o item possui DIF. O segundo tipo de DIF é denominado *DIF não uniforme ou inconsistente*, e se observa quando as CCI's do item estudado, com respeito aos grupos de referência e focal, são diferentes e, ademais, se cruzam em algum ponto do intervalo atitudinal. Em outras palavras, quando há uma vantagem relativa para um desses grupos investigados, cujo valor é variável ao largo de todo o intervalo atitudinal. Este caso ocorre quando os parâmetros  $a$ ,  $b$  ou  $c$  têm valores distintos nas duas CCI's, isto é, quando as CCI's não são paralelas. Este tipo de DIF está representado na Figura 5.

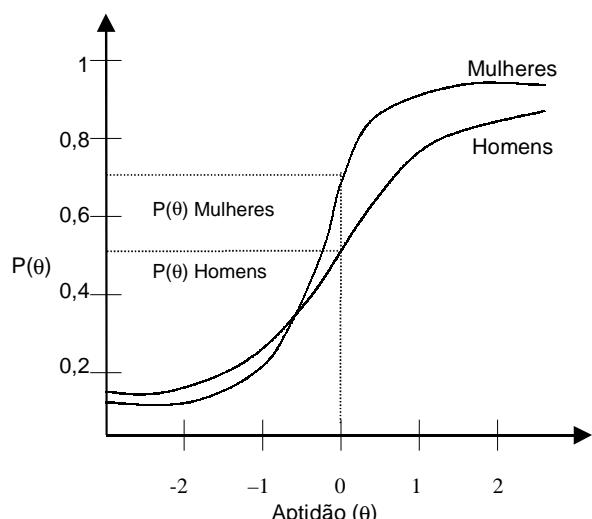


Figura 5. Representação gráfica de um item com DIF não uniforme

A Figura 5 ilustra o caso de diferenças nos parâmetros  $a$ ,  $b$  e  $c$  para os dois grupos investigados. É necessário dizer que neste segundo tipo de DIF é inapropriado um exame global dos dados, dado que tal procedimento poderia ocultar sua presença. É que a natureza variável

Swaminathan & Rogers, 1991  
subjacente à detecção do DIF

- Estimar os parâmetros grupos de comparação;
  - Colocá-los em uma m;
  - Representá-los através d (CCI's);
  - Comparar os valores d os grupos escolhidos.

Já dissemos que, segundo terá DIF se sua CCI não é comparados, cujos sujeitos tem na variável latente  $\theta$  (Kim et al., 2007). A consequência desta definição é adequado para avaliar o DIF entre ambos grupos e sua posterior análise, que difiram significativamente em termos de  $\theta$ , o item tem DIF (Mazor, Hamburgo, & Kellman, 2005).

Não obstante, o problema da investigação psicológica e educacional é que não se consegue com precisão a discrepância entre os resultados obtidos para os distintos subgrupos. Assim, existem muitos procedimentos para o estudo da personalidade, como Holland e Wainer (1993), entrando

- O método das áreas;
  - O método das probabilidades;
  - O método de comparar os critérios métricos dos itens;
  - O método da regressão;
  - O método do *Qui-quadrado*;
  - O método de Mantel-Haenszel.

As investigações para a d  
estão baseadas em uma mesma  
do DIF é um fator que  
interpretação, realizada a parti  
sujeito num item ou teste. É q  
pontuação, seja no âmbito psic  
toda a credibilidade e rep

## Método

### Participantes

Foi composta por 730 estudantes do Ensino Médio, cuja idade média foi 17,7 anos ( $dp = 3,12$  anos), sendo a maioria originária de escolas públicas (58,5%) e pertencendo ao sexo feminino (53,2%).

### Instrumento

Foram utilizadas 30 analogias verbais componentes de um banco de itens já calibrados anteriormente para o uso com a população de estudantes do segundo grau, através do modelo logístico de dois parâmetros (Andriola, 1998).

### Procedimento

Depois dos primeiros contatos com os dirigentes das escolas, para organizar os horários e as turmas que seriam utilizadas no estudo, os 30 itens foram aplicados de modo coletivo na amostra de estudantes. Não houve limitação de tempo para a resolução dos itens.

## Resultados

Para a análise estatística dos dados utilizou-se o programa BILOG-MG, desenvolvido por Zimowski, Muraki, Mislevy e Bock (1996). Este programa utiliza o método de comparação dos parâmetros métricos dos itens para verificar a existência de DIF, mais especificamente a dificuldade do item (parâmetro  $b$ ). Todavia, os itens utilizados no estudo foram calibrados através do modelo TRI de dois parâmetros logísticos, ou seja, tiveram determinados sua discriminação (parâmetro  $a$ ) e dificuldade (parâmetro  $b$ ). Nesse caso, o programa assume como iguais os valores do parâmetro  $a$  para os grupos de comparação (homens e mulheres).

Como afirma Martínez Arias (1997), o uso deste procedimento pretende testar as seguintes hipóteses:

Onde:

- $s\Delta_b$  é o erro-padrão da diferença entre os grupos focal e de referência;
  - $s^2_{bR}$  é a variância do parâmetro  $b$  da referência;
  - $s^2_{bF}$  é a variância do parâmetro  $b$  do grupo focal.
- Neste âmbito, a prova de contraste é realizada a partir da diferença dos parâmetros  $b$  pelo teste estatístico:

$$Z = \frac{\Delta b}{s_{\Delta b}}$$

Onde:

- $Z$  é o resultado da prova de contraste;
- $\Delta b$  é a diferença entre os parâmetros  $b$  do grupo focal e de referência;
- $s\Delta_b$  é o erro padrão da diferença entre os grupos focal e de referência.

Como  $Z$  tem uma distribuição aproximadamente normal, podem ser usadas as tabelas da  $t$ -distribuição para comprovar a significância do valor obtido para a comparação desse valor com o valor crítico.

No nosso caso, o procedimento iniciado para a detecção de DIF consistiu na determinação dos parâmetros métricos dos itens para a amostra total, obtido para o teste da bondade de ajuste (método de máxima verossimilhança) com valor 25.633,96. Em seguida, os parâmetros  $b$  foram determinados para os grupos de comparação (homens e mulheres). Deve ser dito que tais parâmetros devem estar em uma mesma escala. O teste da bondade de ajuste do modelo (método de máxima verossimilhança) com os dados, considerando o grupo de referência, resultou no valor 25.441,77.

A diferença entre os dois valores resultantes da bondade de ajuste do modelo tem significância estatística, como *qui-quadrado* e, se é significativa, indica a existência de DIF no conjunto de itens. A diferença entre o valor inicial (25.633,96) e o valor final resultou em 192,19. Dissemos que esse resultado é estatisticamente significativo, porque

Tabela 1. Valores do Parâmetro  $b$  (dificuldade) dos 30 itens

Itens	Dificuldade dos itens (Parâmetro $b$ )		Diferença entre os grupos (GF-GR)	Err
	Homens(GR)	Mulheres (GF)		
1	3,469	3,303	-0,166	
2	1,132	0,988	-0,145	
3	0,004	-1,401	-1,405	
4	-2,113	-2,327	-0,214	
5	-1,818	-1,460	0,358	
6	-1,264	-1,373	-0,109	
7	-1,302	-1,625	-0,323	
8	-0,844	-1,064	-0,220	
9	-1,209	-1,023	0,186	
10	-0,600	-0,145	0,455	
11	-1,066	-0,913	0,153	
12	-0,877	-1,625	-0,748	
13	-0,210	0,035	0,245	
14	-0,349	-0,021	0,328	
15	-1,101	-0,550	0,551	
16	-1,264	-1,105	0,159	
17	-0,827	-0,483	0,344	
18	-0,600	-0,927	-0,327	
19	-0,287	0,428	0,716	
20	-0,537	-0,671	-0,134	
21	-1,014	-1,091	-0,077	
22	-1,283	-1,358	-0,075	
23	-0,088	-0,430	-0,342	
24	-0,272	-0,281	-0,010	
25	-0,134	0,134	0,268	
26	-0,011	-0,159	-0,147	
27	0,019	0,368	0,349	
28	-0,241	-0,254	-0,013	
29	0,126	0,309	0,183	
30	-0,011	0,149	0,160	

Podemos observar inicialmente que 16 itens (53,3%) apresentam diferenças entre os parâmetros  $b$  dos grupos de homens e mulheres.

### Considerações Finais

O estudo relatado não tinha o objetivo de identificar as causas do DIF, mas verificar sua existência entre 30 analogias componentes de um banco de itens destinados à avaliação do raciocínio verbal em estudantes do ensino médio. Como afirmam Cohen, Kim e Baker (1993), se trata de uma investigação para detectar o DIF, isto é, determinar a possível diferença entre as CCI's dos itens, de acordo com os grupos comparados, adotando, para tanto, o critério de comparação do parâmetro  $b$  (dificuldade) dos itens.

Assim mesmo, como modo de verificar a plausibilidade de se tratar de itens multidimensionais que, supostamente, é uma causa de DIF (Andriola, no prelo b), observamos as cargas fatoriais destes cinco itens no fator único extraído através de análise fatorial (método de máxima verossimilitude). Ditas cargas se situaram entre 0,328 e 0,582, ou seja, não podemos afirmar que tais itens sejam multidimensionais, dado o elevado valor de suas saturações no fator extraído. Não obstante, haveria que tentar identificar outras possíveis causas do DIF destes cinco itens, adotando outros procedimentos, tais como, a análise de conteúdo de ditos itens por experts na área.

É necessário afirmar que a principal contribuição desta investigação para a área da avaliação psicológica e educativa foi identificar os itens com DIF, que são componentes de um banco já organizado e pronto para ser utilizado na avaliação psicológica de estudantes do ensino médio. Os resultados possibilitarão, desse modo, que estes cinco itens não sejam utilizados nos processos de avaliação do raciocínio verbal em dita população de estudantes. Finalmente, a modo de conclusão, queremos apresentar algumas palavras de Pasquali (2000), destacando a importância dos estudos sobre o DIF:

“[...] essas informações sobre cada item de um teste [...] lhe permitem 1) tomar decisões sobre a qualidade do mesmo e 2) colocá-lo num banco de itens sem que ele perca a sua identidade, porque você tem uma série de indicadores

- Camilli, G. & Penfield, D. A. (1997). Variance estimation test functioning based on Mantel-Haenszel. *Educational Measurement, 34*, 123-139.
- Clauser, B. E., Nungester, R. J. & Swaminathan, H. (1993). Item matching for DIF analysis by conditioning on educational background variable. *Journal of Educational Measurement, 33*, 453-464.
- Cohen, A. S., Kim, S. & Baker, F. B. (1993). Detecting differential functioning in the graded response model. *Journal of Educational Measurement, 17*, 335-350.
- Douglas, J. A., Roussos, L. A. & Stout, W. (1996). Item testing: identifying suspect bundles and assessing their functioning. *Journal of Educational Measurement, 33*, 357-375.
- Downing, S. M. & Haladyna, T. M. (1997). Test item analysis: evidence from quality assurance procedures. *Journal of Educational Measurement, 10*, 61-82.
- Hambleton, R. K. (1989a). Introduction. *International Journal of Educational Research, 13*, 123-125.
- Hambleton, R. K. (1989b). Principles and selected topics in test theory. Em R. L. Linn (Org.), *Educational measurement: A handbook for teachers and users* (pp. 147-200). New York: MacMillan.
- Hambleton, R. K. (1994). Item response theory: A unified framework for measurement advances. *Psychological Methods, 19*, 245-270.
- Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). *Principles and applications of item response theory*. North Caroline: Sage.
- Holland, P. W. & Wainer, H. (1993). *Differential item functioning*. Lawrence Erlbaum.
- Jiang, H. & Stout, W. (1998). Improved type I error rate estimation bias for DIF detection using SIBTEST. *Journal of Educational Statistics, 23*, 291-322.
- Kim, S. & Cohen, A. S. (1998). Detection of differential item functioning under the graded response model with the weighted logit link function. *Applied Psychological Measurement, 22*, 345-355.
- Martínez Arias, R. (1997). *Psicometría: Teoría de los tests y sus aplicaciones*. Madrid: Síntesis.
- Mazor, K. H., Hambleton, R. K. & Clauser, B. E. (1993). Effects of item matching on item response theory (IRT) DIF analyses: The effects of matching on unbiased DIF scores. *Applied Psychological Measurement, 22*, 357-375.
- Mellenbergh, G. J. (1989). Item bias and item response theory. *Journal of Educational Research, 13*, 127-143.
- Mislevy, R. J. (1996). Test theory reconceived. *Journal of Educational Measurement, 33*, 379-416.
- Muñiz, J. (1997). *Introducción a la teoría de respuesta a los ítems*. Pirámide.
- Muñiz, J. & Hambleton, R. K. (1992). Medio siglo de la teoría de respuesta a los ítems. *Anuario de Psicología, 52*, 41-66.
- Oshima, T. C., Raju, N. S. & Flowers, C. P. (1997). Demonstration of multidimensional IRT-based differential functioning of items and tests. *Journal of Educational Measurement, 34*, 253-272.

- Williams, V. S. L. (1997). The “unbiased” anchor: Bridging the gap between DIF and item bias. *Applied Measurement in Education, 10*, 253-267.
- Zimowski, M. F., Muraki, E., Mislevy, R. J. & Bock, R. D. (1996). *BILOG-MG. Multiple Group IRT Analysis and Test Maintenance for Binary Items*. Chicago: Scientific Software International (SSI).

Sobre o autor

**Wagner Bandeira Andriola** é Professor do Mestrado em Avaliação Educacional da Universidade Federal do Ceará (UFC); Psicólogo pela Universidade Federal da Paraíba (UFPB); Especialista em Psicometria pela Universidade de Brasília (UnB); Mestre em Psicologia Social e do Trabalho pela Universidade de Brasília (UnB); Doutorando do Programa *Investigación, Diagnóstico y Evaluación para la Calidad Educativa da Universidad Complutense de Madrid (UCM)*; Bolsista da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Atua nas áreas de Instrumentação e Medida Psicológica e Educativa utilizando as Teorias Clássica dos Testes (TCT) e de Resposta ao Item (TRI), com especial interesse pelo estudo do funcionamento diferencial dos itens (DIF) em testes psicológicos e educativos.



**Objetivo Geral:** Implementar a formação do Desenvolvimento Humano de profissionais, técnicos, agentes de comunidade e estudantes das áreas de Educação e Saúde, através de discussões teórico-temáticas, dinâmicas pedagógicas e produção de material pedagógico.