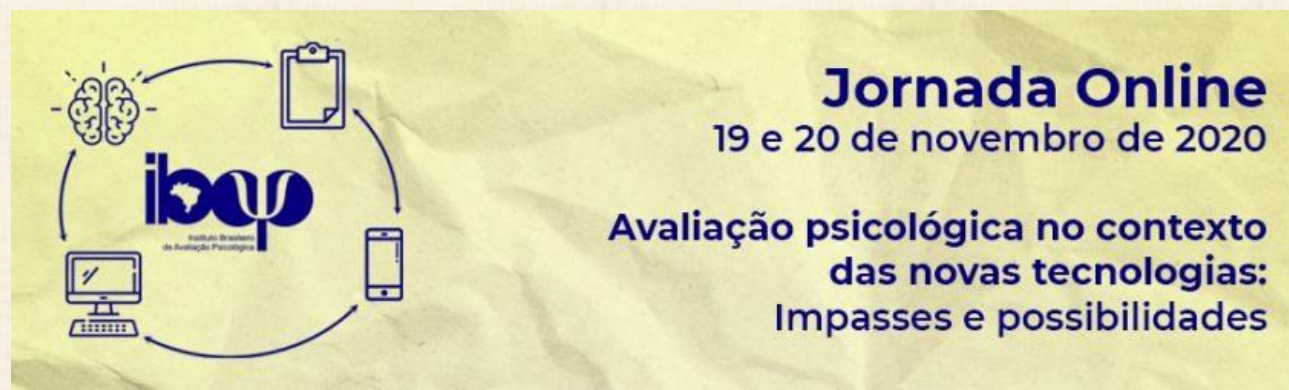


Análise de Equivalência e Invariância de Estrutura de Instrumentos Psicológicos



Dr. Felipe Valentini &
Dr. Nelson Hauck



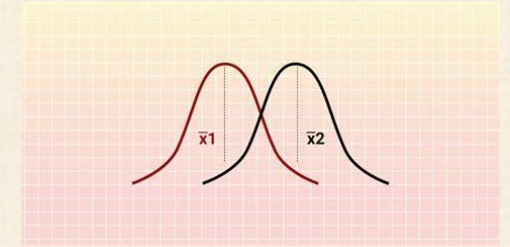
Antes de tudo....



Não...

Não é esta fake news..



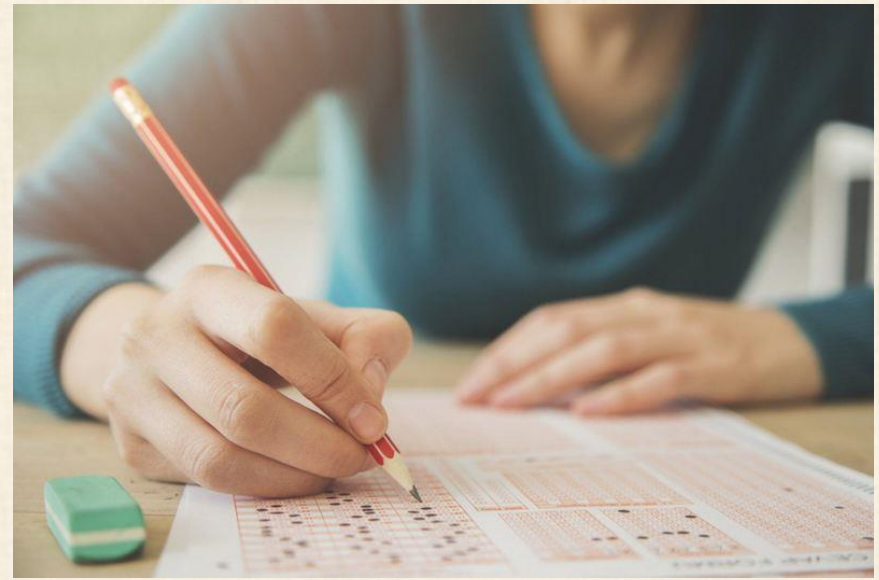


Antes de tudo....

Teste t & Anova não testam equivalência ou invariância...

Esses testes partem do pressuposto da invariância...

Mas vamos chegar lá... Aguardem...



Pensar sempre...

Testagem envolve dois elementos principais:

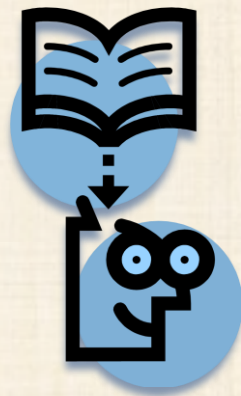
- **Características do examinando**
- **Características dos itens/instrumento**

Aplicações...

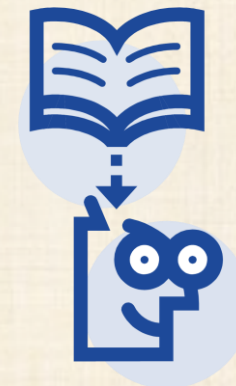


- O teste ‘funciona’ de maneira igual entre dois grupos (ou mais de dois grupos)?
 - ex. teste de atenção mantém a mesma parametrização entre homens e mulheres
- Diferentes formas de apresentar um item produzem escores compatíveis?
 - Ex. teste lápis-e-papel vs informatizado

Marciano
Acertou 18 questões

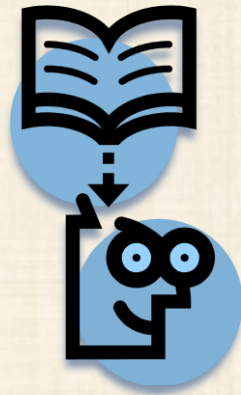


Terráqueo
Acertou 19 questões



Quem é mais inteligente ?

Marciano
Acertou 18 questões
Do teste WAIS



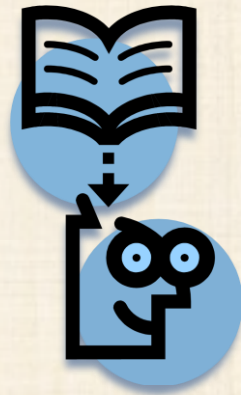
Terráqueo
Acertou 19 questões
Do teste RAVEN



Antes de responder à pergunta é necessário garantir que os testes sejam equivalentes...

Aplico o mesmo teste.... WAIS

Marciano continua
acertando 18 questões
Tem 25 anos



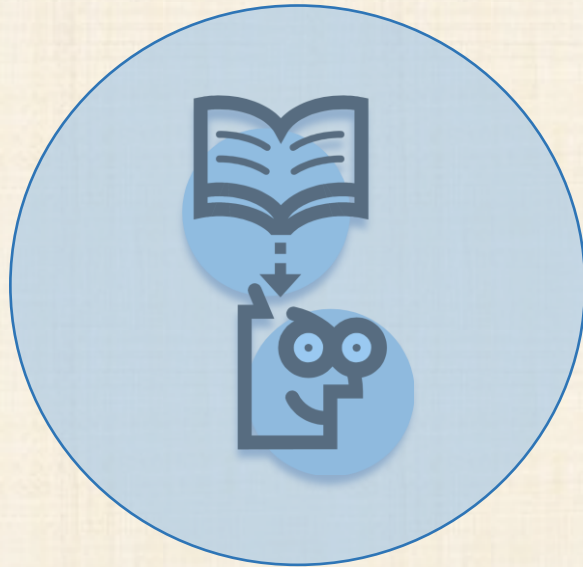
Terráqueo continua
acertando 19 questões
Tem 25 anos



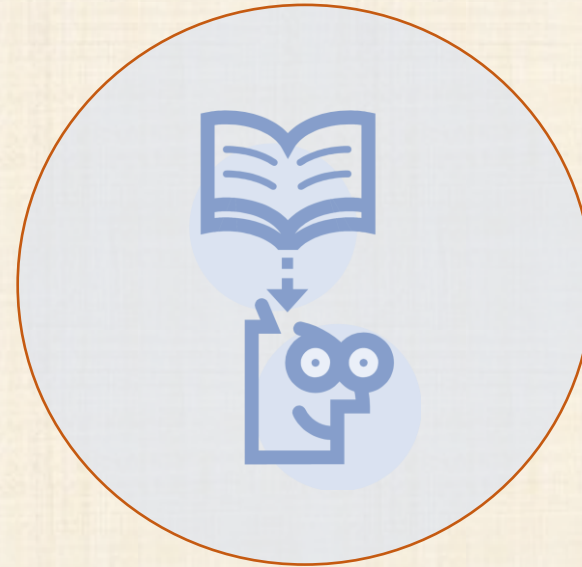
Quem é mais inteligente ???

Aplico o mesmo teste.... WAIS

Marciano é de Marte...



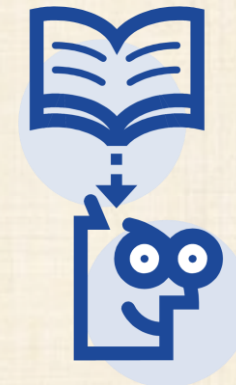
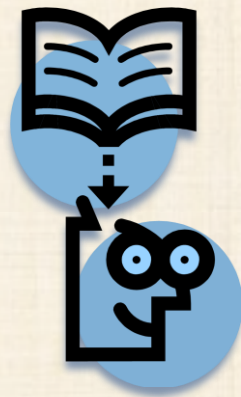
Terráqueo é da Terra...



Antes de responder à pergunta é necessário garantir que os testes (ou os escores) não sejam enviesados...

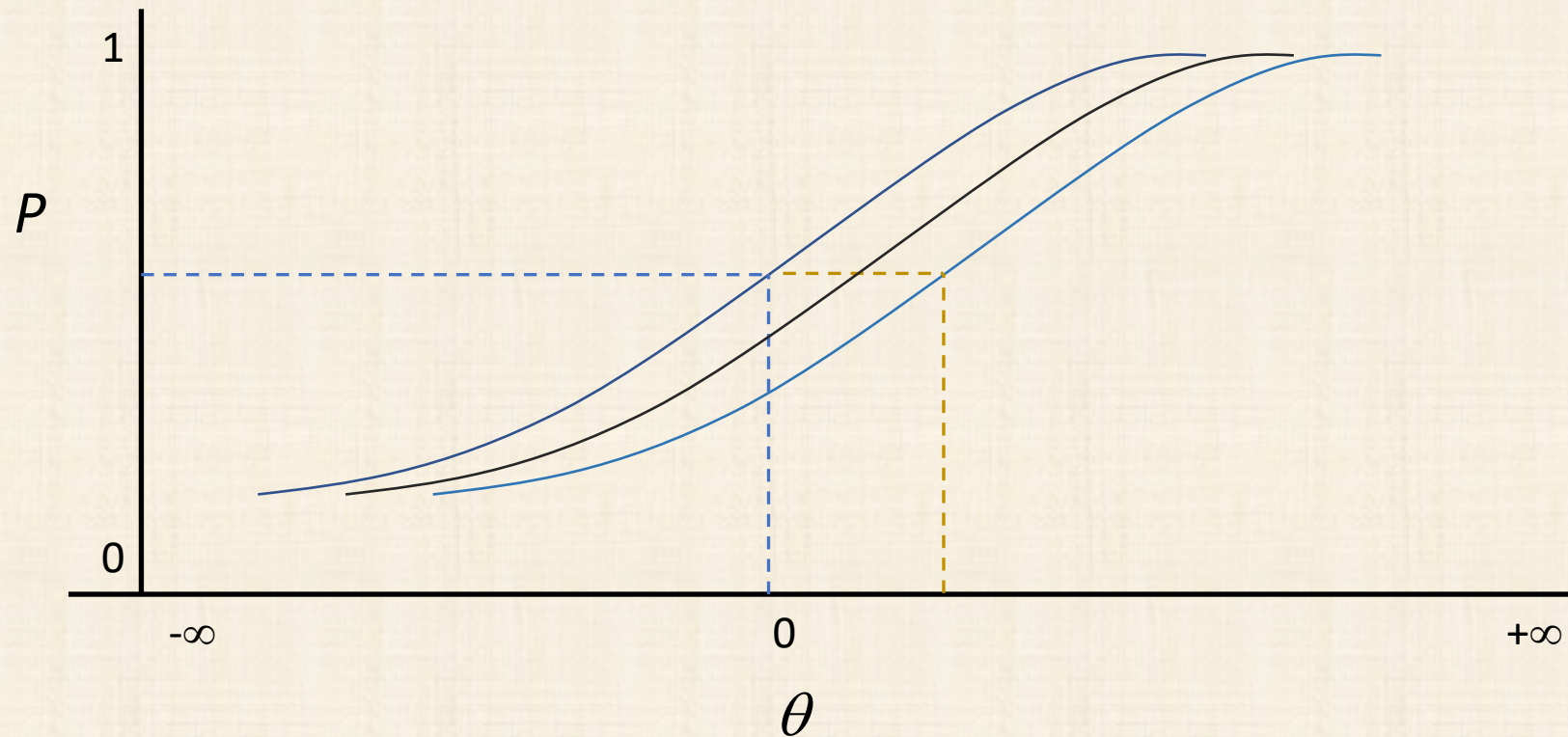
A estrutura do instrumento e os parâmetros variam de população para população ?

Se: $\theta_m = \theta_t$

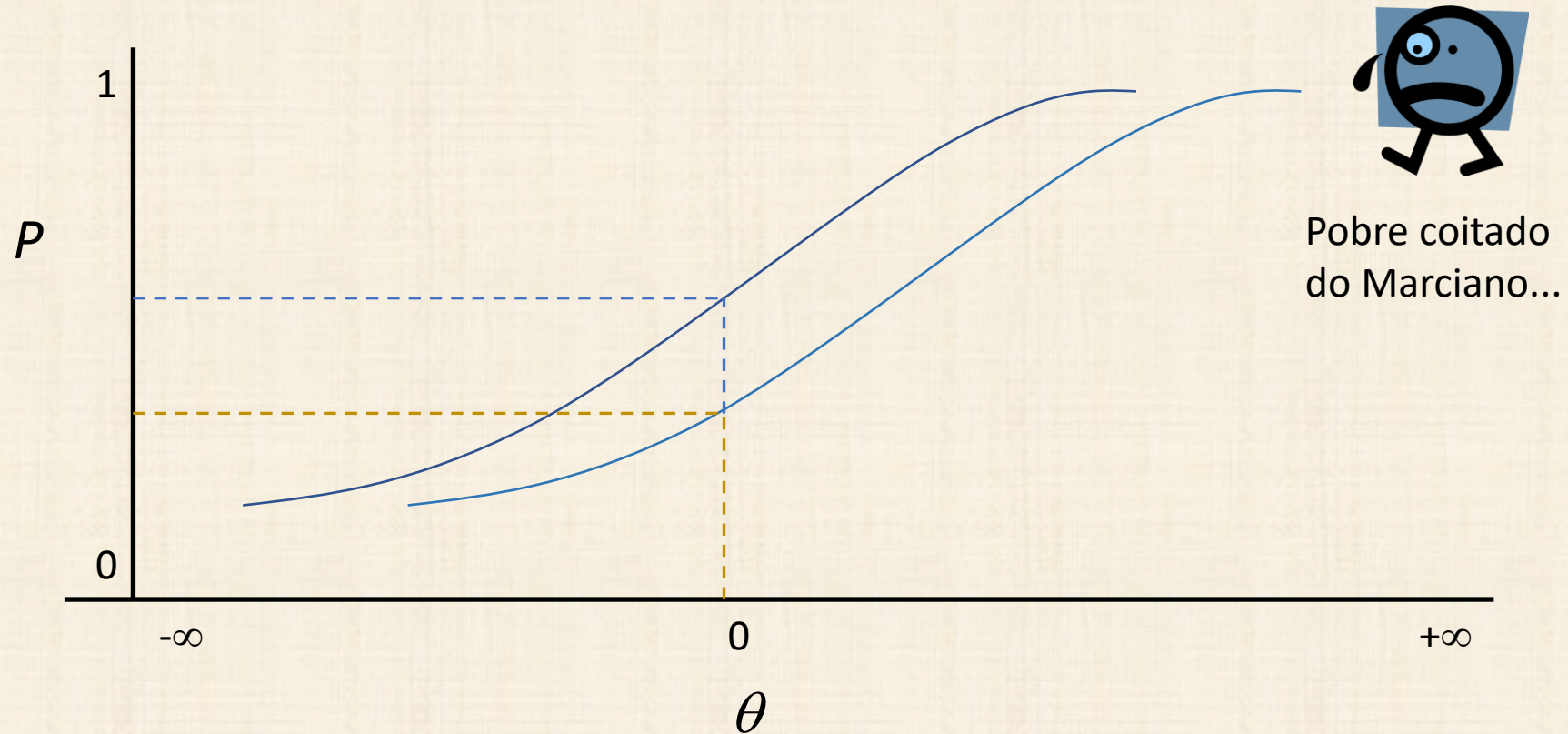


Deve... $P(\text{item } 1=1) = P(\text{item } 1=1)$

A estrutura do instrumento e os parâmetros variam de população para população ?



A estrutura do instrumento e os parâmetros variam de população para população ?



Diferentes técnicas

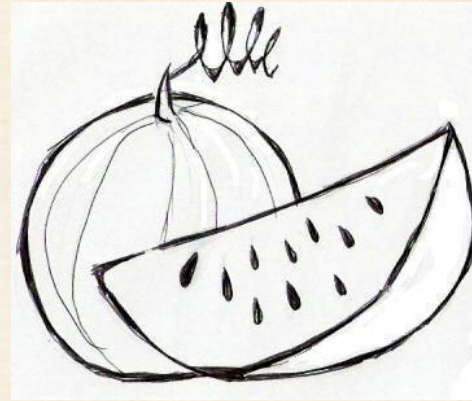
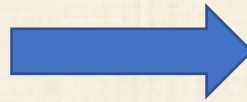
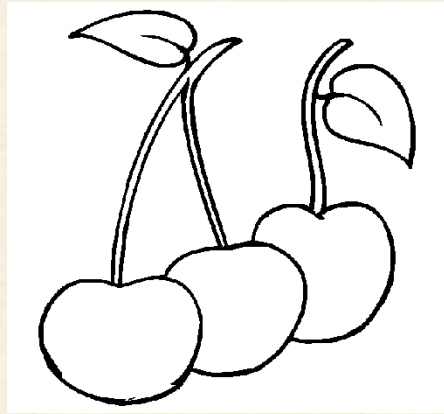
Também denominada de...

Measurement Invariance (Análise de Invariância)

Measurement Equivalence (Equivalência da Medida)

- DIF
- Análise Fatorial Multigrupos
- MIMIC
- Invariância para medidas pareadas

Exemplo Teste SON-R 2 ½ - 7



Análise Fatorial Multigrupos

Aplicação: para dois ou mais grupos.

- Inicia com um modelo o mais livre possível
 - Quase todos os parâmetros dos itens são liberados para variar (serem diferentes) entre os grupos
 - São fixados apenas alguns parâmetros para fins de identificação
 - Modelo inicial (é o utilizado para comparar com os subsequentes)
- Nos modelos seguintes são restritos alguns parâmetros como iguais entre os grupos
 - Se o ajuste do modelo piorar, significa que a restrição dos parâmetros não funcionou, portanto não há invariância.
 - Se o ajuste do modelo não piorar, há invariância...

Lembrem-se da equação de regressão...

Lembre-se: Temos
itens e **pessoas**

$$Y = B_0 + B_1 X + E$$

Vamos embaçar um pouco...

Para indivíduos i e j ; grupos g e h

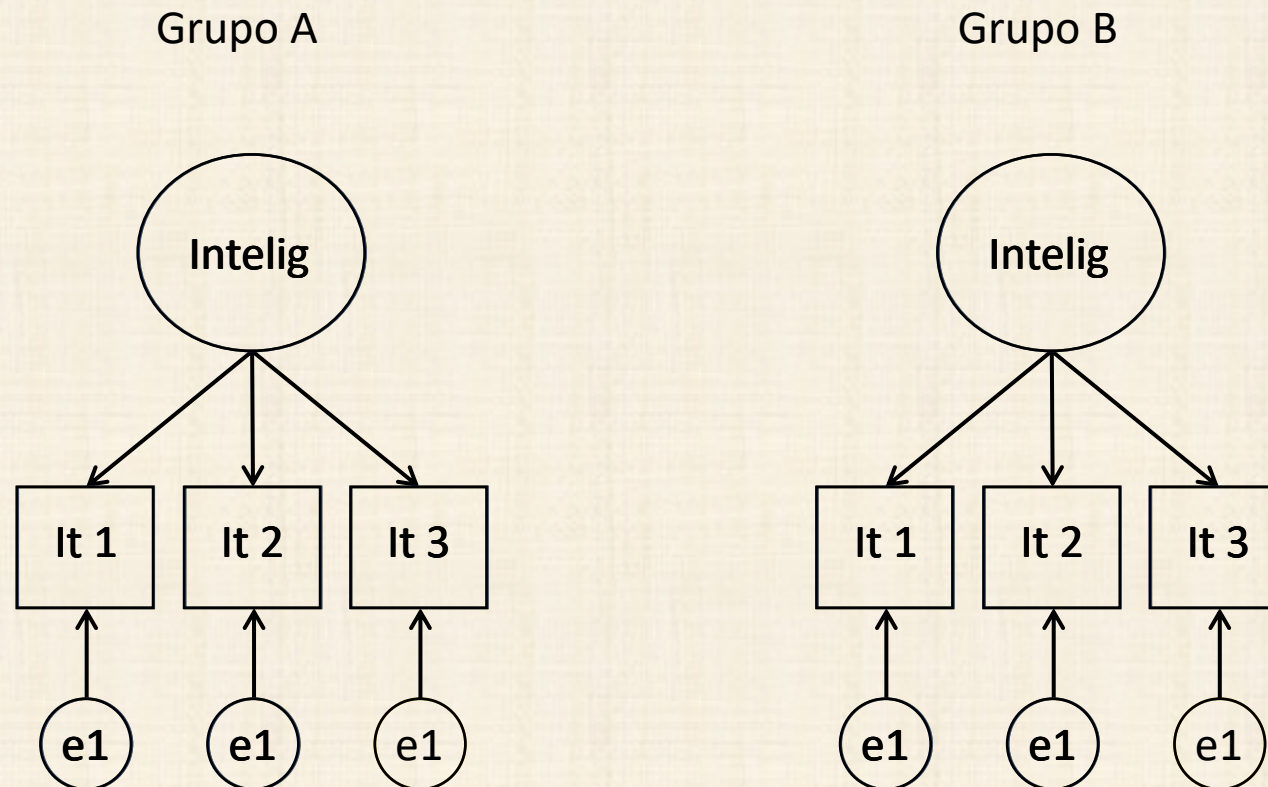
$$\begin{aligned} Y_{ig} &= \alpha_g + \beta_{1g} X_{ig} + \varepsilon_{ig} \\ Y_{jh} &= \alpha_h + \beta_{1h} X_{jh} + \varepsilon_{jh} \end{aligned}$$

Se esses parâmetros
forem fixos, a
diferença de Y será
uma função apenas
da diferença de X

Tipo	Definição (Fixar)
1 - Invariância Configural	Número de itens e fatores
2 – Invariância Métrica (Weak Invariance)	Número de itens e fatores + cargas fatoriais
3 – Invariância Escalar (Strong Invariance)	Número de itens e fatores + cargas fatoriais + Interceptos
4 – Invariância Residual (Strict Invariance)	Número de itens e fatores + cargas fatoriais + Interceptos + variância residual

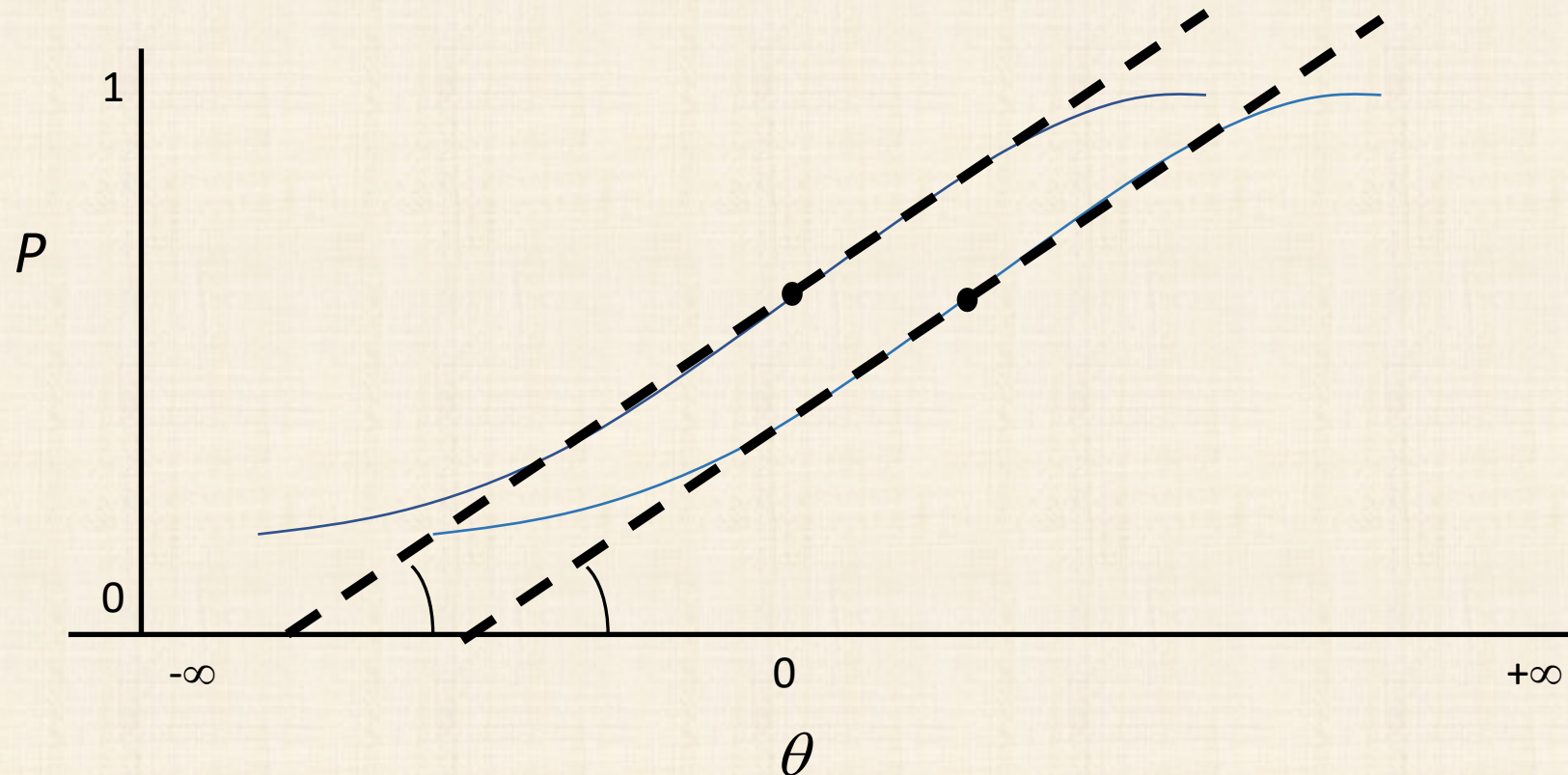
- Invariância Configural

A estrutura do modelo é plausível para todos os grupos?



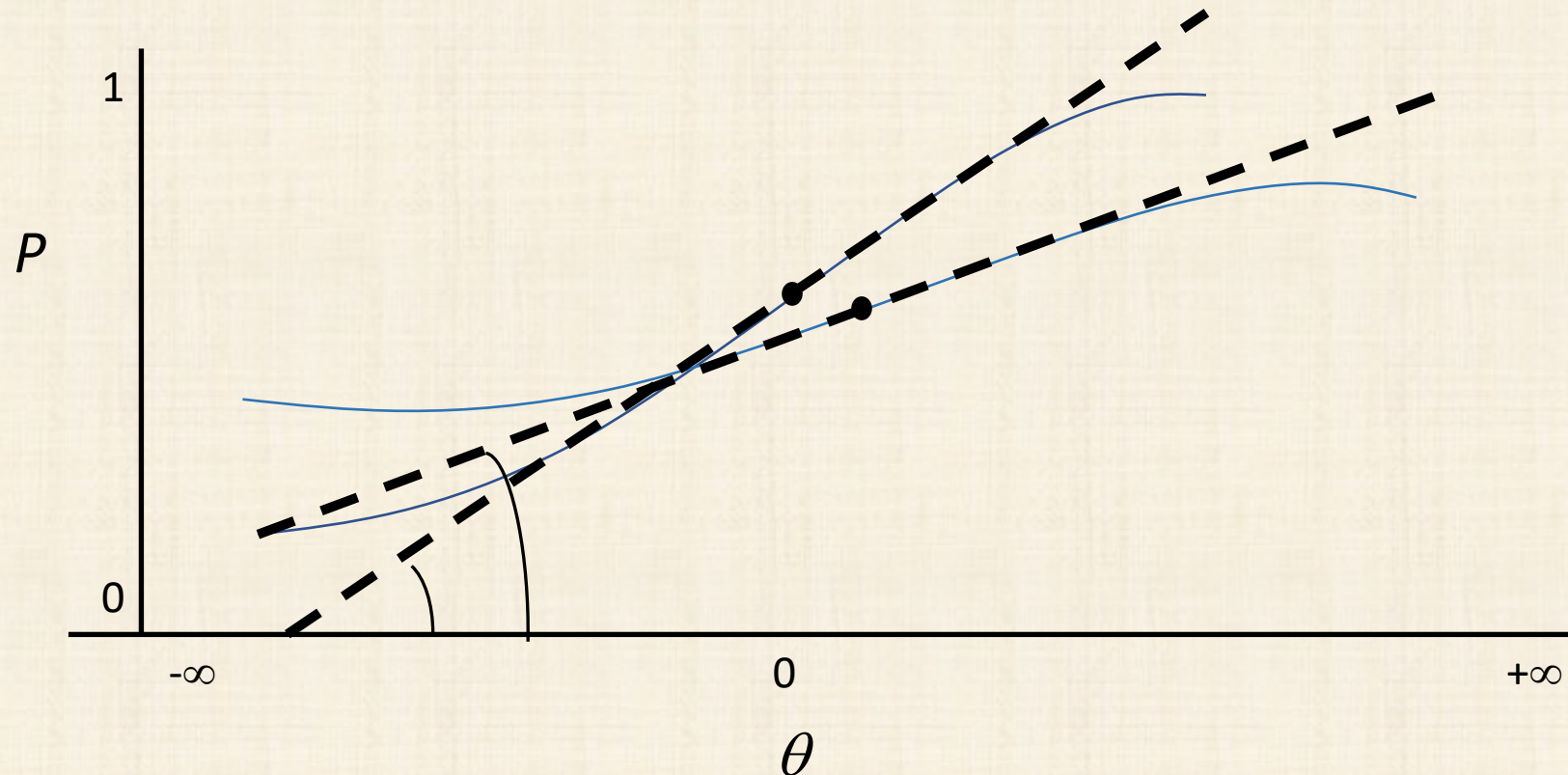
- Invariância Métrica

Além do número de fatores e itens, as cargas fatoriais são invariantes? Os itens têm a mesma importância para os diferentes grupos?



- Invariância Métrica

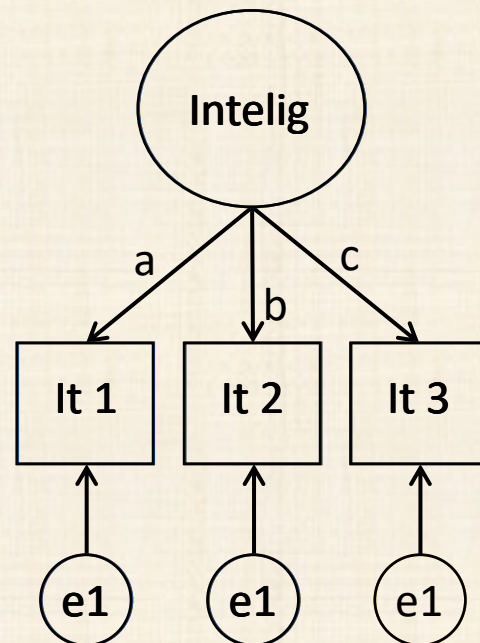
Além do número de fatores e itens, as cargas fatoriais são invariantes? Os itens têm a mesma importância para os diferentes grupos?



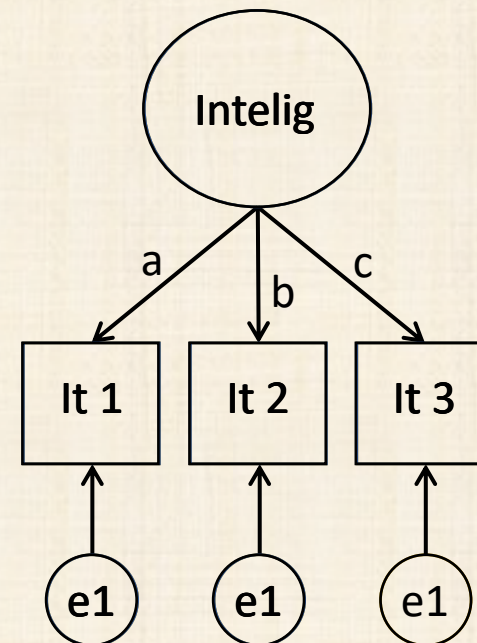
- Invariância Métrica

Além do número de fatores e itens, as cargas fatoriais são invariantes? Os itens têm a mesma importância para os diferentes grupos?

Grupo A



Grupo B



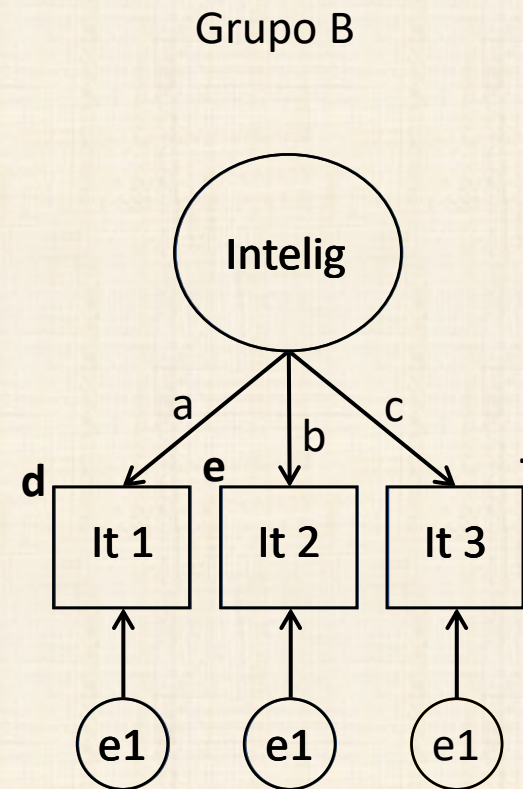
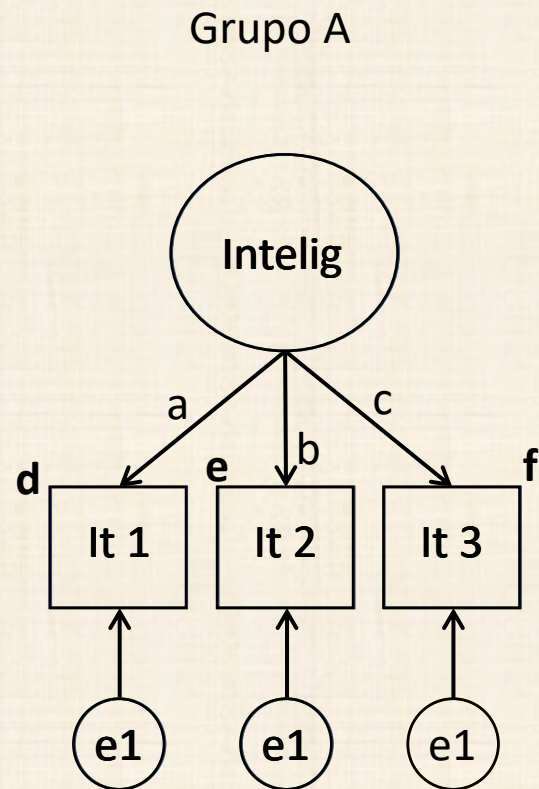
- Invariância Escalar

Além dos passos anteriores, os interceptos são iguais?

Teste está relacionado à dificuldade dos itens/ou probabilidade de endosso...

Portanto, se a invariância dos interceptos não for garantida, eventuais diferenças entre os grupos, podem ser creditadas à vieses relacionados à dificuldade dos itens e não, necessariamente, às diferenças reais entre os grupos...

- Invariância Escalar
Interceptos fixos...



- Invariância Residual

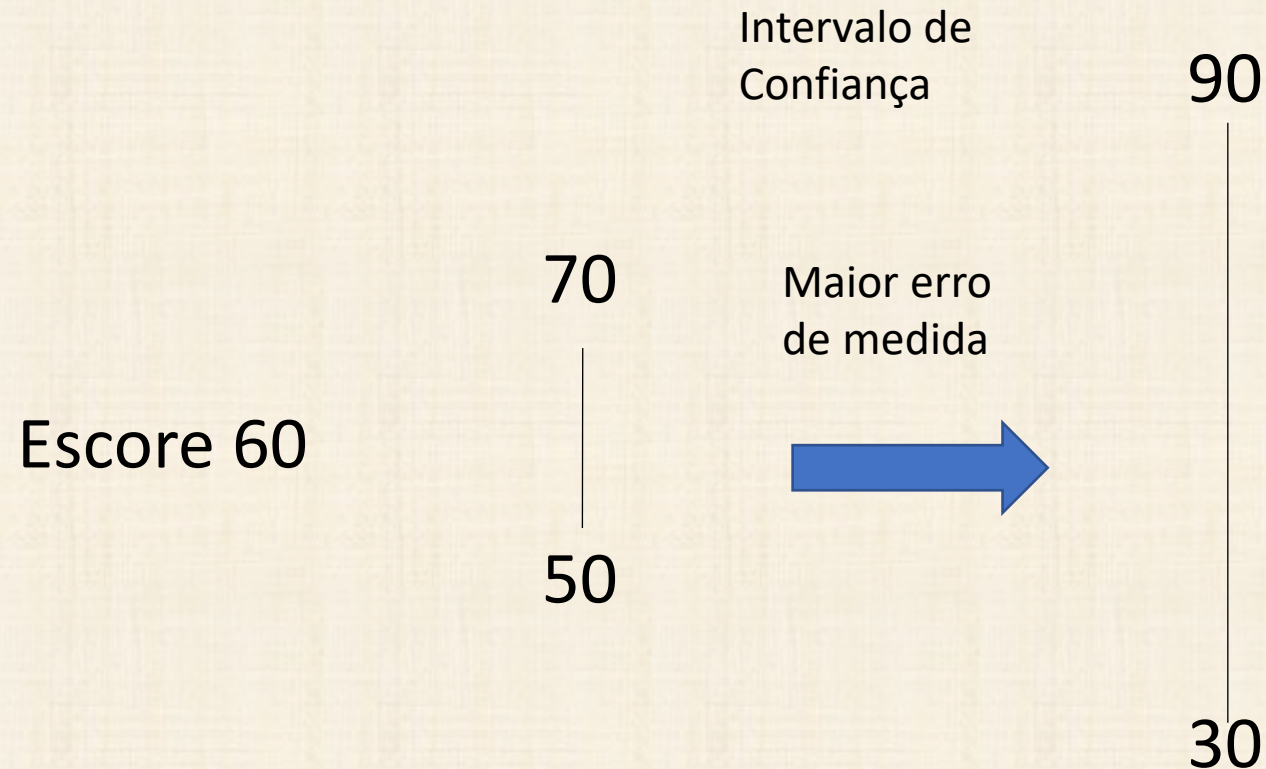
Além dos passos anteriores, os erros/resíduos são equivalentes?

Os resíduos estão associados ao erro de medida = contrário à precisão...

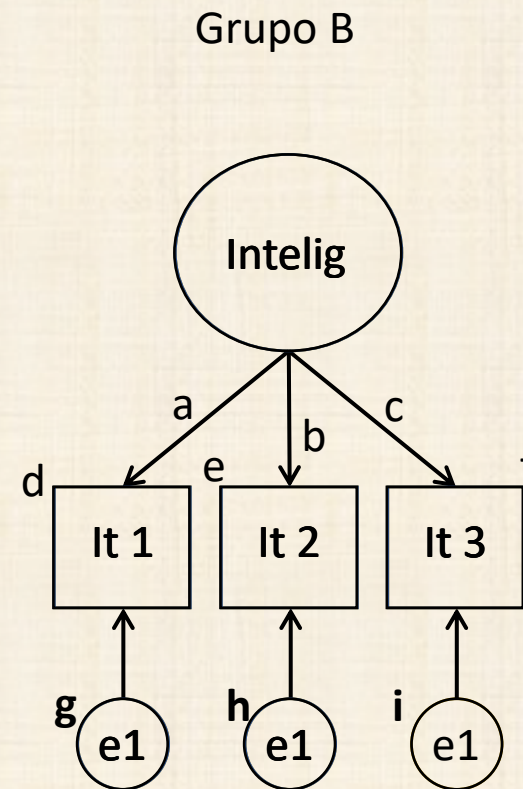
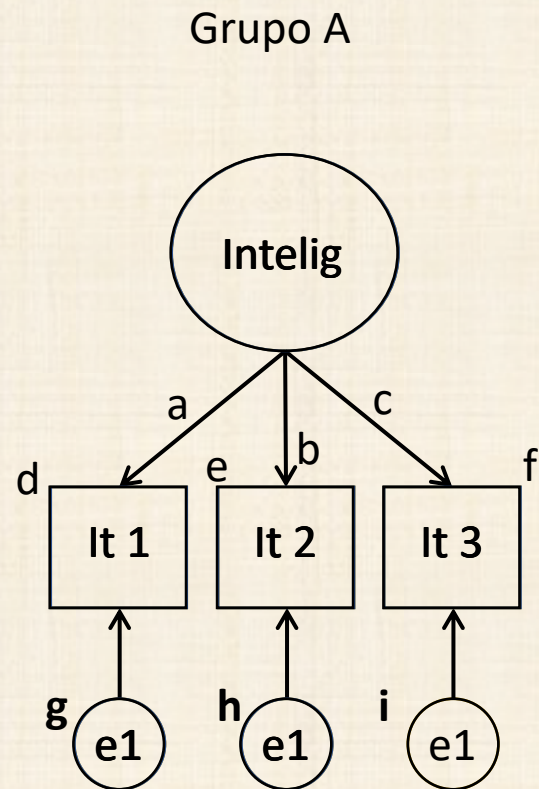
Portanto, a invariância residual diz respeito à manutenção da mesma precisão para a estimação dos escores dos diferentes grupos...

- Invariância Residual

Você já sabe que... Todo o escore pode variar em função da imprecisão ao medir... Se tivermos como estimar a precisão é possível calcular um intervalo de confiança do escore...



- Invariância Residual
Erros fixos...



Como testar?

- Modelos mais restritos, normalmente, apresentam indicadores de ajuste mais singelos/piores do que modelos mais livres...
- Análise de Invariância impõe, sistemática e paulatinamente, restrições ao modelo...
- Portanto, basta fixar os parâmetros e avaliar o ajuste... Se o modelo piorar muito... Significa que não é possível assumir aquele tipo de invariância...

Modelo	χ^2 (gl)	$\Delta\chi^2$ (Δ gl)	CFI	MC	Gamma-hat
Configural	5 (10)	-	0,981		
Métrica	6 (15)	1 (5)	0,980		
Escalar	8 (20)	2 (5)	0,975		
Residual	30 (25)	22 (5)	0,940		

$\Delta\chi^2$ - não significativo

Δ CFI < 0,01

Pontos de corte, ver Svetina & Rutkowski (2019) Multiple-group invariance with categorical outcomes using updates guidelines: an illustration using Mplus and the lavaan/semTools Packages. *Structural Equation Modeling: A multidisciplinary journal*.

TABLE 1

Selected Examples of Varying Cutoff Values for Testing ME/I under Various Approaches

<i>Source</i>	<i>Approach</i>	<i># Groups</i>	<i>N</i>	<i># Factors</i>	<i>Distribution</i>	<i>Overall recommendations*</i>
Chen (2007)	MG-CFA	2	150, 250, or 500 per group	1	Normal	$\Delta CFI \geq -.005$, $\Delta RMSEA \leq .010$ $\Delta CFI \leq -.005$ or $-.010$ for CFI, $\Delta RMSEA \geq .010$ or $.015$ $\Delta \Gamma_{\text{hat}} \leq -.005$ or $-.008$. $\Delta SRMR \geq .025$ or $.030$ for metric invariance testing $\Delta SRMR \geq .005$ or $.010$ for intercept and residual variance invariance testing
Cheung and Rensvold (2002)	MG-CFA	2	150 or 300 per group	2 or 3	Normal	$\Delta CFI \geq -.010$, $\Delta \Gamma_{\text{hat}} \geq -.001$, $\Delta \text{McDonald's NCI} \geq -.02$
French and Finch (2006)	MG-CFA	2	150/150 150/500 or 500/500	2 or 4	Normal	ΔCFI less than $-.01$ or chi-square difference of $p < .05$ or $.01$ (with use of maximum likelihood)
French and Finch (2006)	MG-CFA	2	150/150 150/500 or 500/500	2 or 4	Ordinal	$\Delta \chi^2$ at $.05$ (but low power)
Finch and French (2018)	Equivalence testing	2	100, 200, 400, 600, 1000, 1500, or 2000 per group	1		$\varepsilon 0^+$ (Equivalence) for some value of RMSEA Excellent fit: < 0.01 Close fit: $0.01-0.05$ Fair fit: $0.05-0.08$ Mediocre fit: $0.08-0.10$ Poor fit: $0.10+$
Kim et al. (2017) [†]	MG-CFA ML-CFA ML- FMM Bayesian Alignment	25 or 50	50, 100, or 1000 per group	1	Normal	MG CFA, ΔCFI with the cutoff of $.01$ ML CFA, ΔCFI with the cutoff of $.01$ BIC with total sample size Bayesian, the PPP of $.05$ and 95% CI
Rutkowski and Svetina (2014)	MG-CFA	10 or 20	Varied from 600 to 6,000 per group	1	Normal	$\Delta RMSEA \leq .03$ and $\Delta CFI \geq .020$ for metric; $\Delta RMSEA \leq .01$ and $\Delta CFI \geq .010$ for scalar
Rutkowski and Svetina (2017)	MG-CFA	10 or 20	Varied from 600 to 6,000 per group	1	Ordinal	$\Delta RMSEA \leq .05$ in conjunction with sig. $\Delta \chi^2$ and $\Delta CFI \geq -.004$ for metric $\Delta RMSEA \leq .01$ in conjunction with sig. $\Delta \chi^2$ and $\Delta CFI \geq -.004$ for scalar
Svetina and Rutkowski (2017)	MG-CFA	10 or 20	750 to 6,000 per group	2 or 5	Ordinal	$\Delta RMSEA \leq .05$ in conjunction with significant $\Delta \chi^2$ for metric $\Delta RMSEA \leq .01$ and $\Delta CFI \geq .002$ for scalar [for 3 or fewer dimensions]

- Os modelos estudados até agora partem do pressuposto que os itens são medidas intervalares...
- Para itens categóricos ou ordinais há parâmetros adicionais de Thresholds.

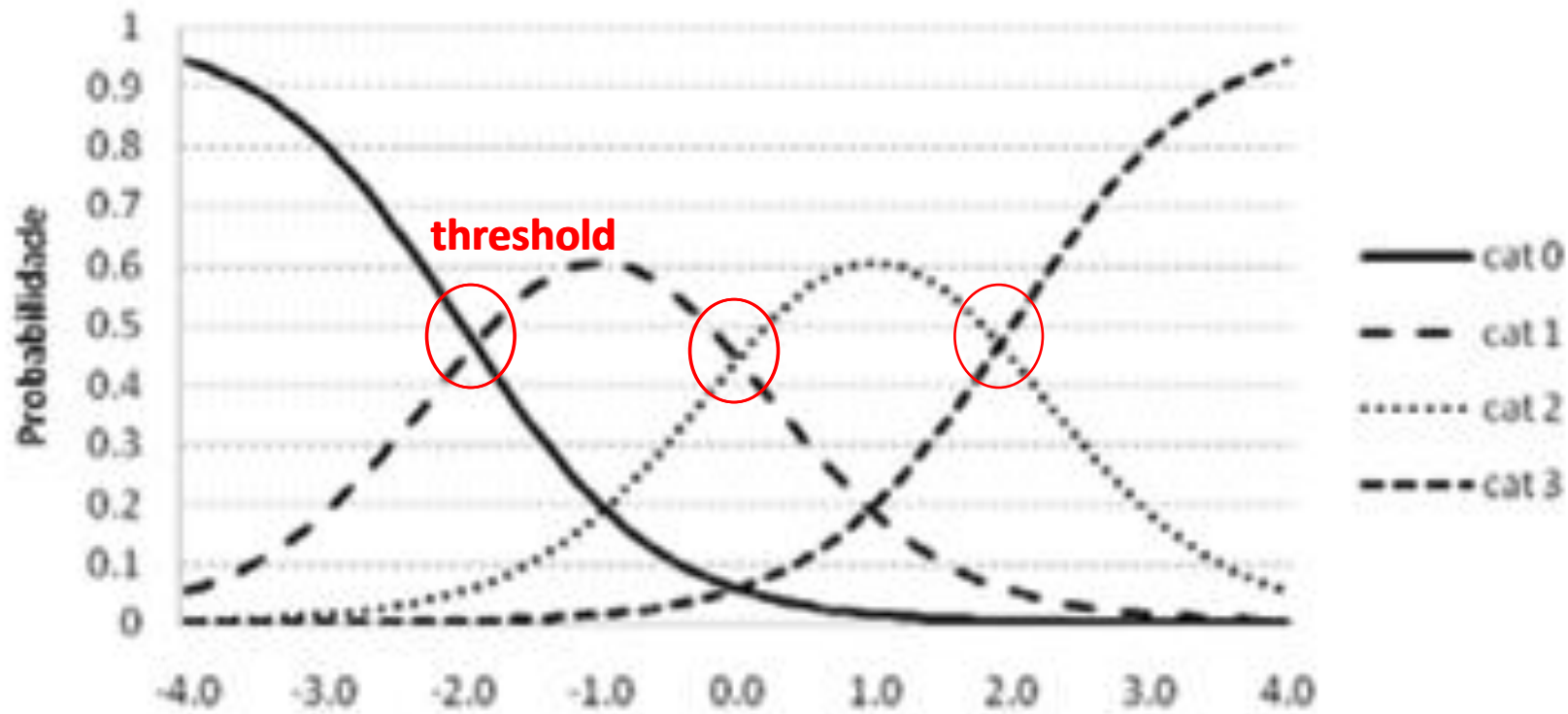



Figura 2 – CCI do Modelo de Resposta Gradual de um item com $a = 1,4$ e $b_1 = -2,0$, $b_2 = 0,0$, $b_3 = 2,0$.

Solução Wu

PSYCHOMETRIKA
DOI: 10.1007/s11336-016-9506-0



IDENTIFICATION OF CONFIRMATORY FACTOR ANALYSIS MODELS OF DIFFERENT LEVELS OF INVARIANCE FOR ORDERED CATEGORICAL OUTCOMES

HAO WU 

BOSTON COLLEGE

RYNE ESTABROOK

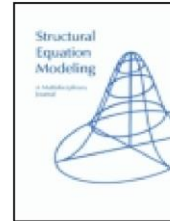
NORTHWESTERN UNIVERSITY

This article considers the identification conditions of confirmatory factor ordered categorical outcomes with invariance of different types of parameters. The practice of invariance testing is to first identify a model with only configural invariance of parameters based on this identified baseline model. This approach identifies different identification conditions on this baseline model identify the scales of 1 in different ways. Once an invariance condition is imposed on a parameter, they may become restrictions and define statistically non-equivalent models, leading to analyzing the transformation that leaves the model-implied probabilities of response identification conditions for models with invariance of different types of parameters to a specific parametrization of the baseline model. Tests based on this approach they do not depend on the specific identification condition chosen for the baseline model.

Key words: ordered categorical data, invariance testing, model identification.

1. Introduction

Ordered categorical measures such as Likert scales are prevalent. Data collected with such measures had been analyzed as normal data until ordered categorical outcomes were introduced (Jöreskog & Moustaki, 2001, 1984) and proved to be superior (Babakus, Ferguson & Jöreskog, 1989; Lubke & Muthén, 2004; Muthén & Kaplan, 1985). One of the challenges for ordered categorical outcomes is the threshold model (see, e.g., Muthén, 1985). ordered categories are considered a discretized version of a normally distributed continuous variable.



Structural Equation Modeling: A Multidisciplinary Journal



ISSN: 1070-5511 (Print) 1532-8007 (Online) Journal homepage: <https://www.tandfonline.com/loi/hsem20>

Multiple-Group Invariance with Categorical Outcomes Using Updated Guidelines: An Illustration Using *Mplus* and the lavaan/semTools Packages

Dubravka Svetina, Leslie Rutkowski & David Rutkowski

To cite this article: Dubravka Svetina, Leslie Rutkowski & David Rutkowski (2019): Multiple-Group Invariance with Categorical Outcomes Using Updated Guidelines: An Illustration Using *Mplus* and the lavaan/semTools Packages, Structural Equation Modeling: A Multidisciplinary Journal, DOI: 10.1080/10705511.2019.1602776

To link to this article: <https://doi.org/10.1080/10705511.2019.1602776>



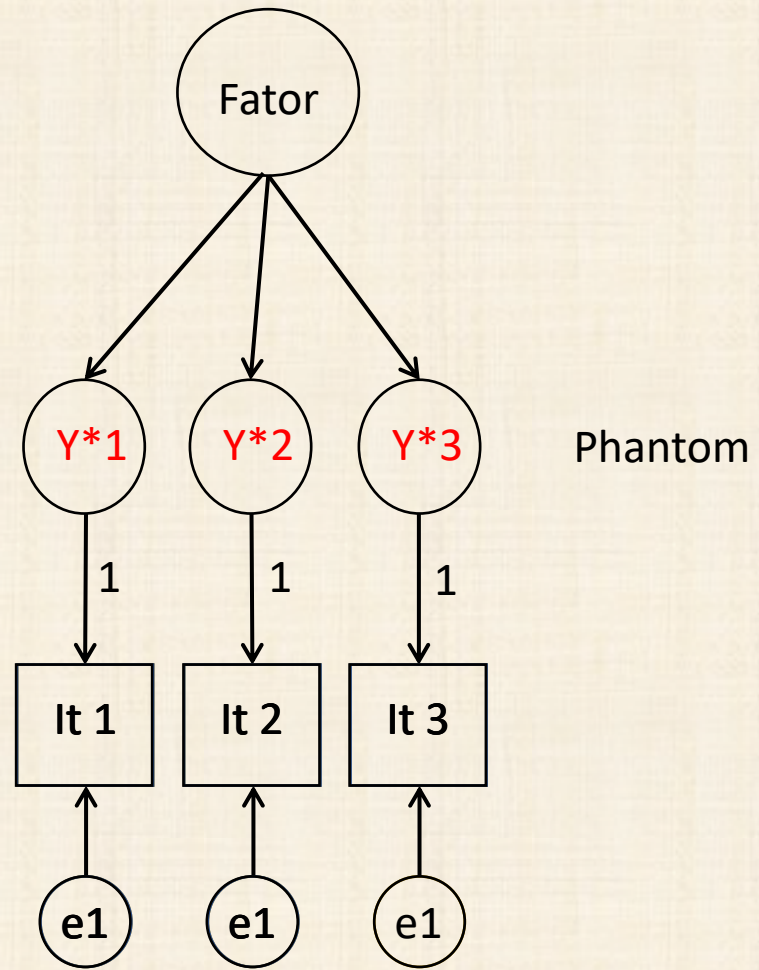
Published online: 29 Apr 2019.



Submit your article to this journal [↗](#)



Article views: 59



Solução WU

- Sugestão
- Baseline
 - Máximo livre, mantendo parâmetros fixos somente aqueles necessários para identificação
- Proposição 4
 - Restringir os thresholds como iguais entre os grupos
 - Cargas e interceptos livres
- Proposição 7
 - Restringir as cargas
 - Interceptos livres
- Modelo final
 - Restringir os interceptos
 - Liberar as médias dos fatores

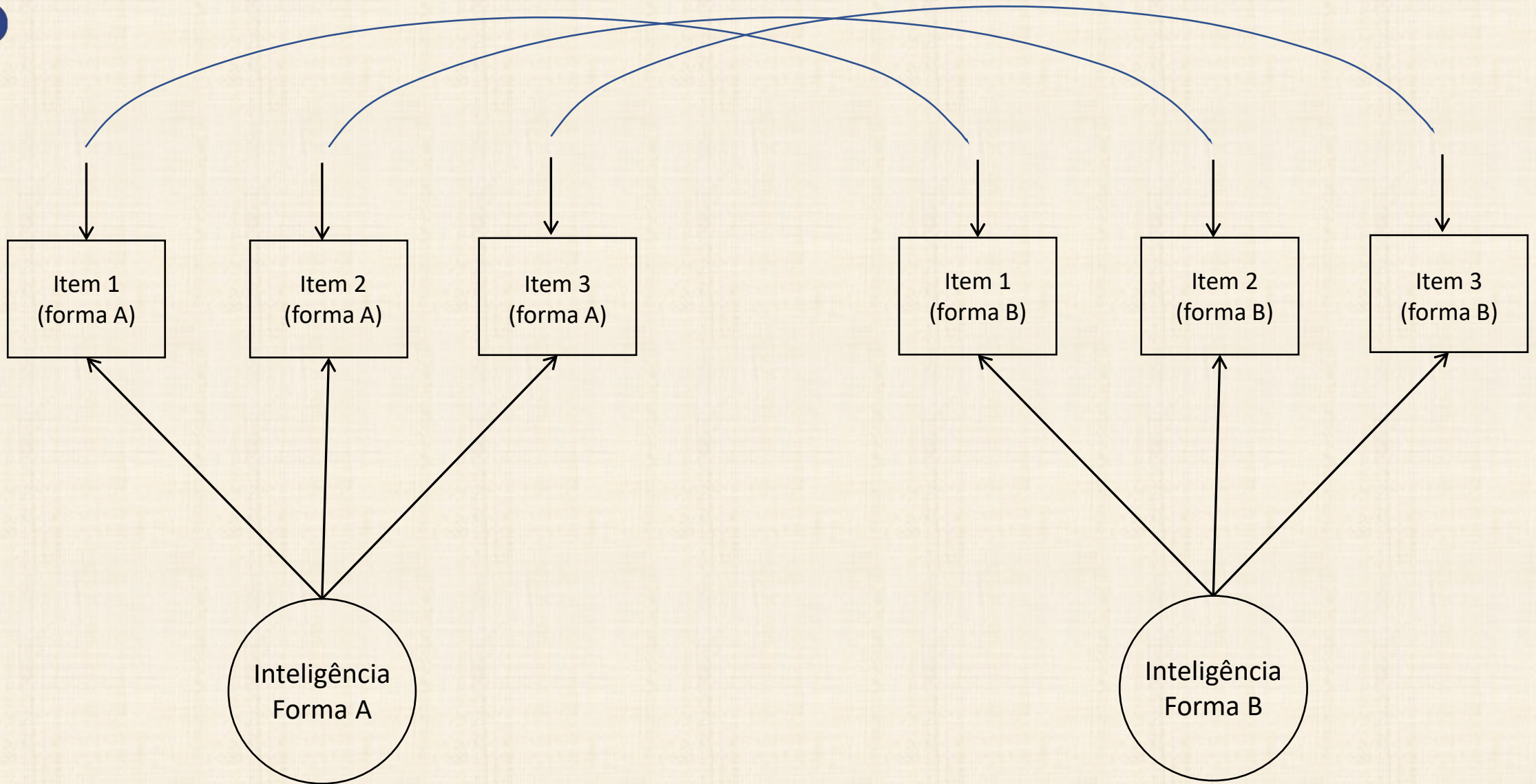
	Baseline	Prop4 (thresholds restritos)	Prop 7 (thresholds + cargas restritas)	Prop 7 + interceptos
cargas	livres	livres	restritos entre grupos	restritos entre grupos
interceptos (phanton no Mplus)	fixo 0 (todos grupos)	fixo 0 primeiro grupo + livre nos demais grupos	fixo 0 primeiro grupo + livre nos demais grupos (parece um erro do tutorial)	fixo 0 (todos os grupos)
variância phantons	Fixa em 0 (todos os grupos)	Fixa em 0 (todos os grupos)	fixo 0	fixo 0
Thresholds	Livres	restritos entre grupos	restritos entre grupos	restritos entre grupos
Médias latentes	fixo 0 (todos grupos)	fixo 0 (todos grupos)	fixo 0 (todos grupos)	fixo 0 primeiro grupo + livre nos demais grupos
Variância Latente	fixo 1 (todos os grupos)	fixo 1 (todos os grupos)	fixo 1 primeiro grupo + livre nos demais grupos	fixo 1 primeiro grupo + livre nos demais grupos
Scale	fixo 1 (todos os grupos)	fixo 1 primeiro grupo + livre nos demais grupos	fixo 1 primeiro grupo + livre nos demais grupos	fixo 1 primeiro grupo + livre nos demais grupos
Variância Residual	livres (1 - carga ²)	livres (1 - carga ²)	livres (1 - carga ²)	livres (1 - carga ²)
Lavaan chama de variância dos itens				

Diferentes técnicas

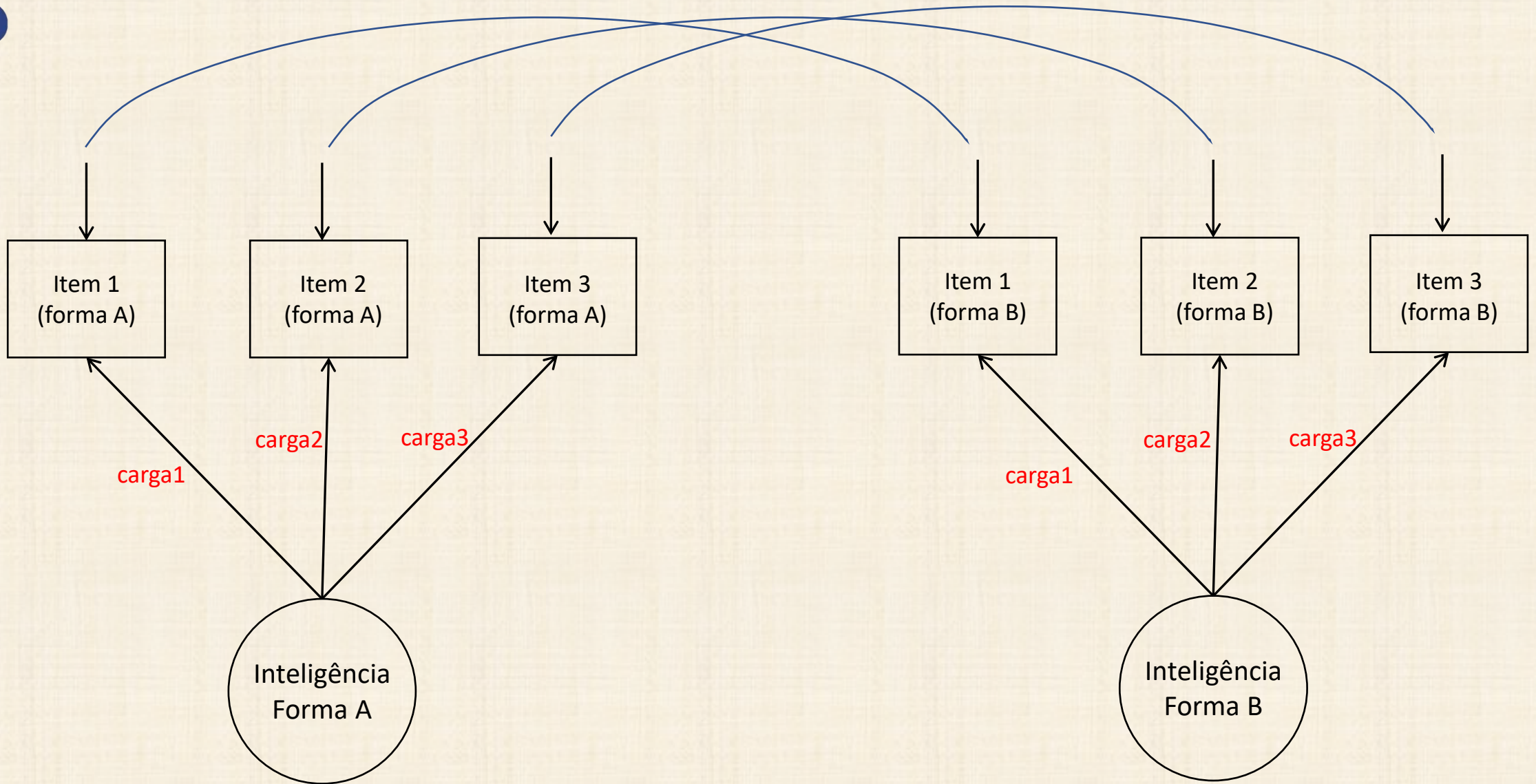
- DIF
- Análise Fatorial Multigrupos
- MIMIC
- Invariância para medidas pareadas

Diferentes técnicas

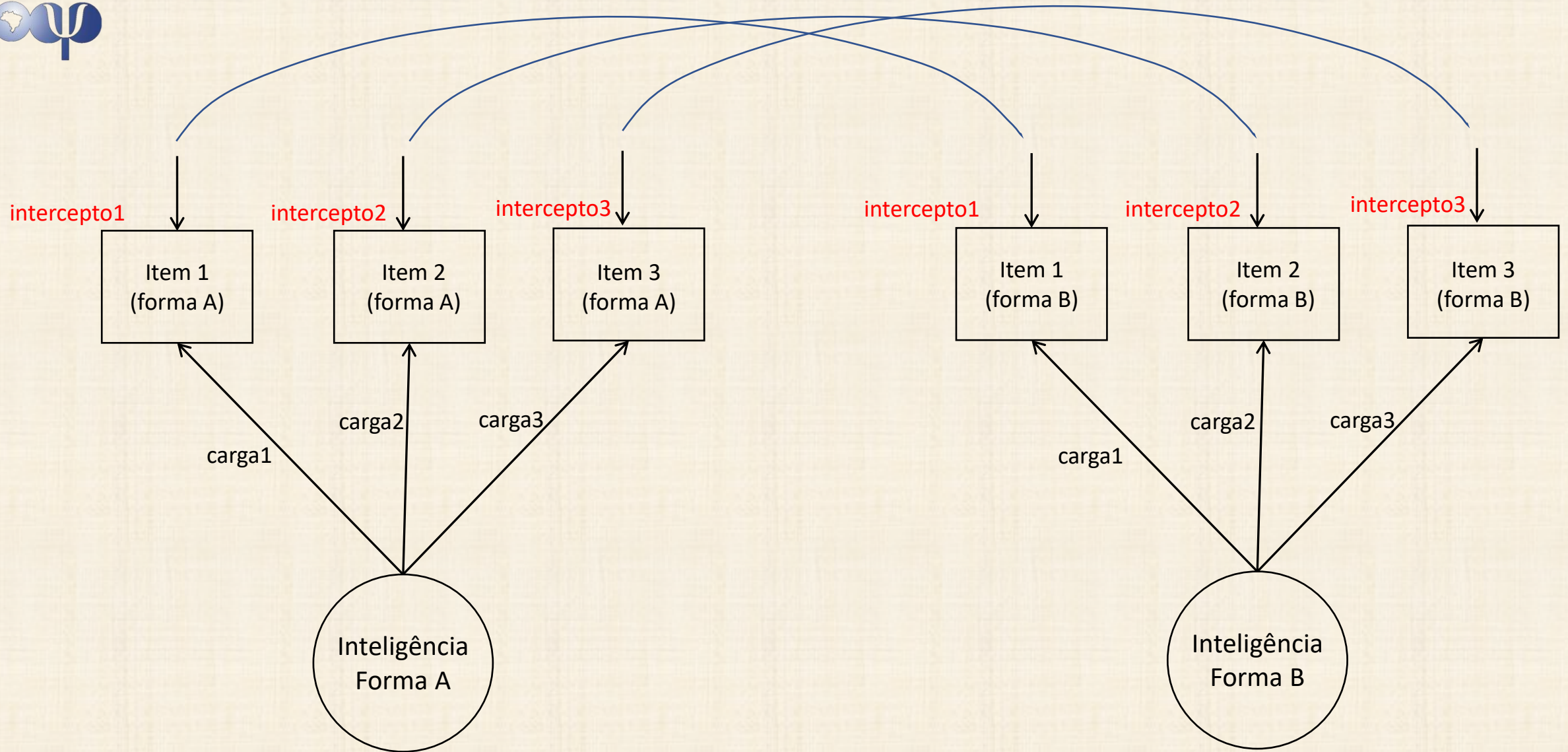
- **Invariância para medidas pareadas**
 - Quando as mesmas pessoas respondem a versões diferentes do teste
 - Ex. lápis e papel vs informatizado
 - Raven no tempo 1 e repete o mesmo teste 6 meses depois...



Modelo baseline (configural)



Fixar CARGAS como iguais entre as formas (medidas pareadas)



Fixar INTERCEPTOS como iguais entre as formas (medidas pareadas)