



Análise de itens e pressupostos da análise fatorial exploratória (AFE)

Prof. Dr. Wagner de Lara Machado



Análise de itens e pressupostos da análise fatorial exploratória (AFE)

A AFE é uma das técnicas mais utilizadas na psicologia, especialmente na **psicometria**, com a finalidade de testar o desempenho de instrumentos psicométricos ou **gerar teorias** (Costello & Osborne, 2005; Haig, 2005)

MULTIVARIATE BEHAVIORAL RESEARCH, 40(3), 303–329
Copyright © 2005, Lawrence Erlbaum Associates, Inc.

Exploratory Factor Analysis, Theory Generation, and Scientific Method

Brian D. Haig
University of Canterbury

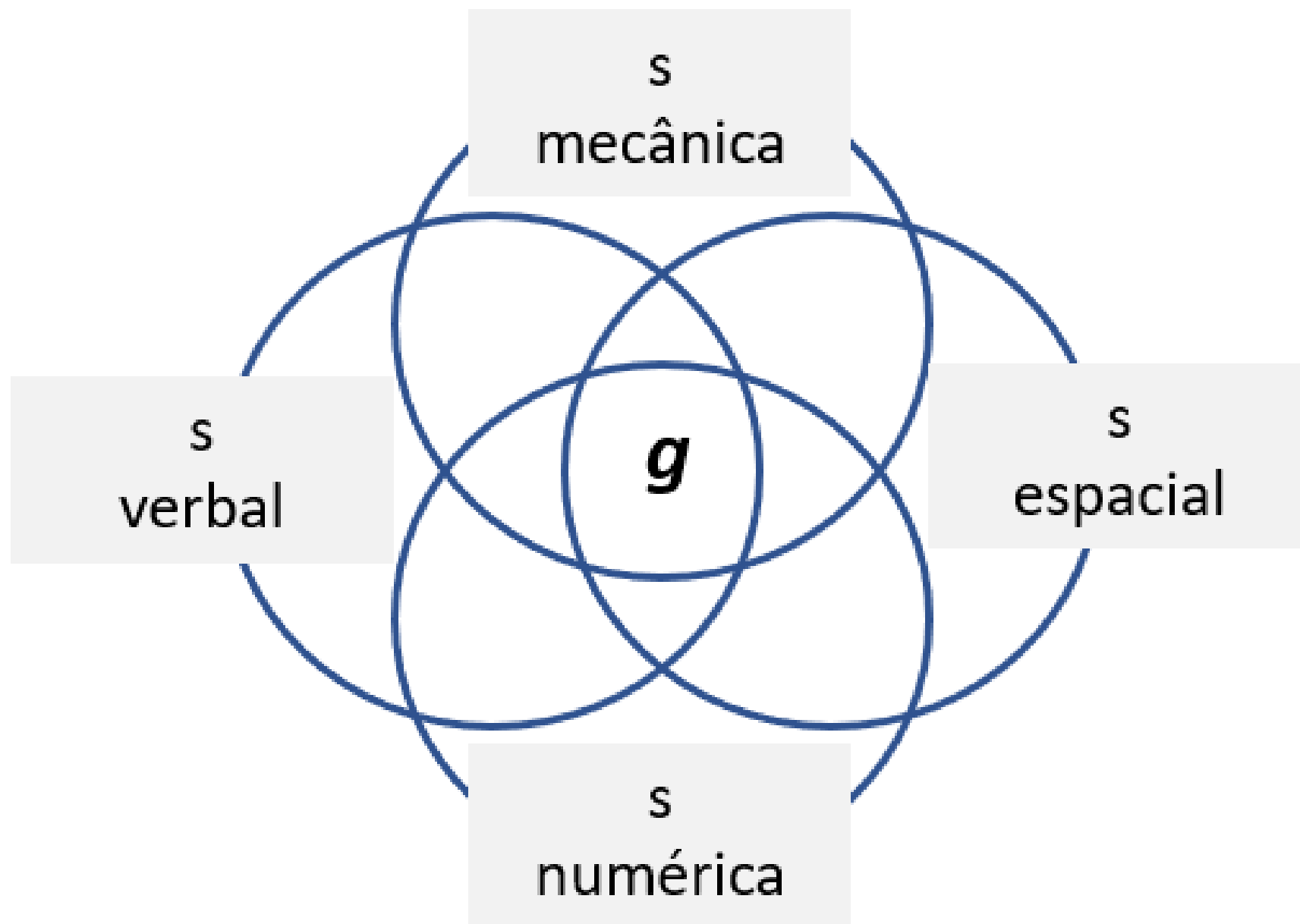
Charles Spearman



Foi desenvolvida por Spearman e utilizada na geração da teoria do fator geral de inteligência (g)



Sua hipótese era de que havia uma variável (geral) que explicava as relações entre habilidades específicas



Cinco grandes fatores da personalidade (big five)

Por Que Cinco Fatores?

A descoberta dos cinco fatores foi acidental e se constitui em uma generalização empírica, replicada independentemente inúmeras vezes. Como o modelo não foi desenvolvido a partir de uma teoria, não há, consequentemente, uma explicação teórica *a priori* (e satisfatória) dos motivos que levariam a organização da personalidade em cinco (e não quatro, ou sete) dimensões básicas.

Psicología Reflexão e Crítica
Universidade Federal do Rio Grande do Sul
prcrev@ufrgs.br
ISSN: 0102-7972
BRASIL

1998

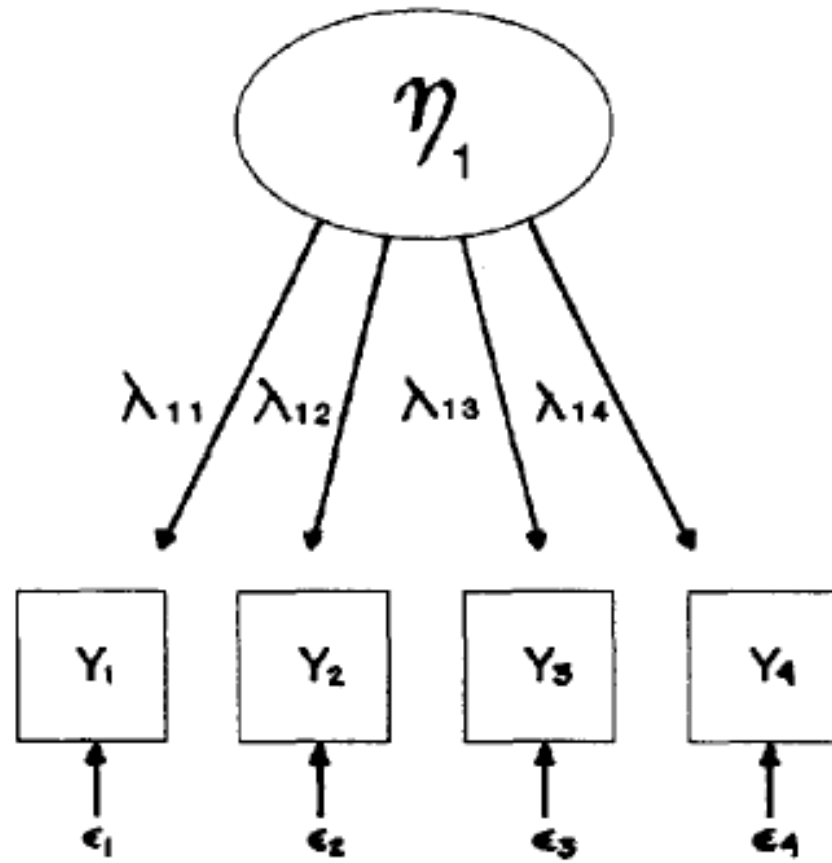
Claudio Simon Hutz / Carlos H. Nunes / Alice D. Silveira / Jovana Serra / Márcia Anton
/ Luciane S. Wieczorek

O DESENVOLVIMENTO DE MARCADORES PARA A AVALIAÇÃO DA
PERSONALIDADE NO MODELO DOS CINCO GRANDES FATORES



Psicometria

- Na Psicometria, o modelo de AFE é empregado para descobrir (!) o número de variáveis (**latentes**) necessárias e suficientes para explicar um conjunto de itens ou tarefas de um instrumento psicométrico. Um uso bem comum é na investigação das propriedades psicométricas de escalas, questionários e inventários



USO DA ANÁLISE FATORIAL EXPLORATÓRIA EM PSICOLOGIA

Bruno Figueiredo Damásio¹ – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil

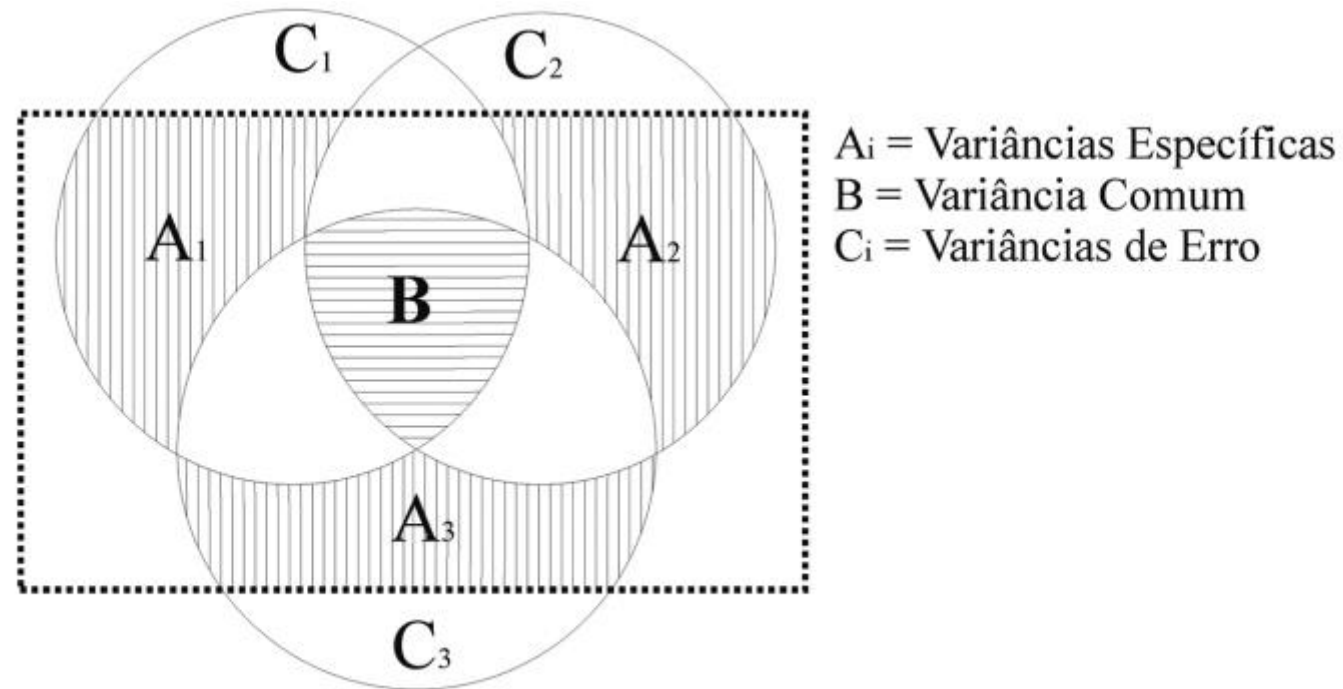


Figura 1 - Ilustração das variâncias de três itens e suas relações com um fator hipotético.

Em suma

- A AFE é uma técnica que
- 1) parte da hipótese de que a correlação entre variáveis tem como **causa** uma ou mais variáveis, também chamada de “princípio da causa comum”;
- 2) introduz a modelagem de **variáveis latentes**, isto é, variáveis que não são diretamente observáveis, e, por isso, não estão no banco de dados;
- 3) diferente de outras técnicas estatísticas, a AF é exploratória e interpretativa, alinhada ao raciocínio abutivo (geração de hipóteses e teorias após a observação empírica)

Formulação matemática

- $X = b + a_1F_1 + a_2F_1 + a_3F_1 \dots + e$

- Em que:

X é o escore observado no teste;

b é um intercepto, em geral fixado em zero;

a é o coeficiente angular, indicando a relação linear entre item e fator;

F é um escore fatorial, valor latente do traço em questão;

e é um erro aleatório com média zero e distribuição normal.

Pressupostos

- Distribucionais – método de extração
- Correlações – qualidade dos dados

Anexo

Escala de Satisfação com a Vida

Instruções

Abaixo você encontrará cinco afirmações com as quais pode ou não concordar. Usando a escala de resposta a seguir, que vai de 1 a 7, indique o quanto concorda ou discorda com cada uma; escreva um número no espaço ao lado da afirmação, segundo sua opinião. Por favor, seja o mais sincero possível nas suas respostas.

7 = Concordo totalmente

6 = Concordo

5 = Concordo ligeiramente

4 = Nem concordo nem discordo

3 = Discordo ligeiramente

2 = Discordo

1 = Discordo totalmente

1. ____ Na maioria dos aspectos, minha vida é próxima ao meu ideal.

2. ____ As condições da minha vida são excelentes.

3. ____ Estou satisfeito(a) com minha vida.

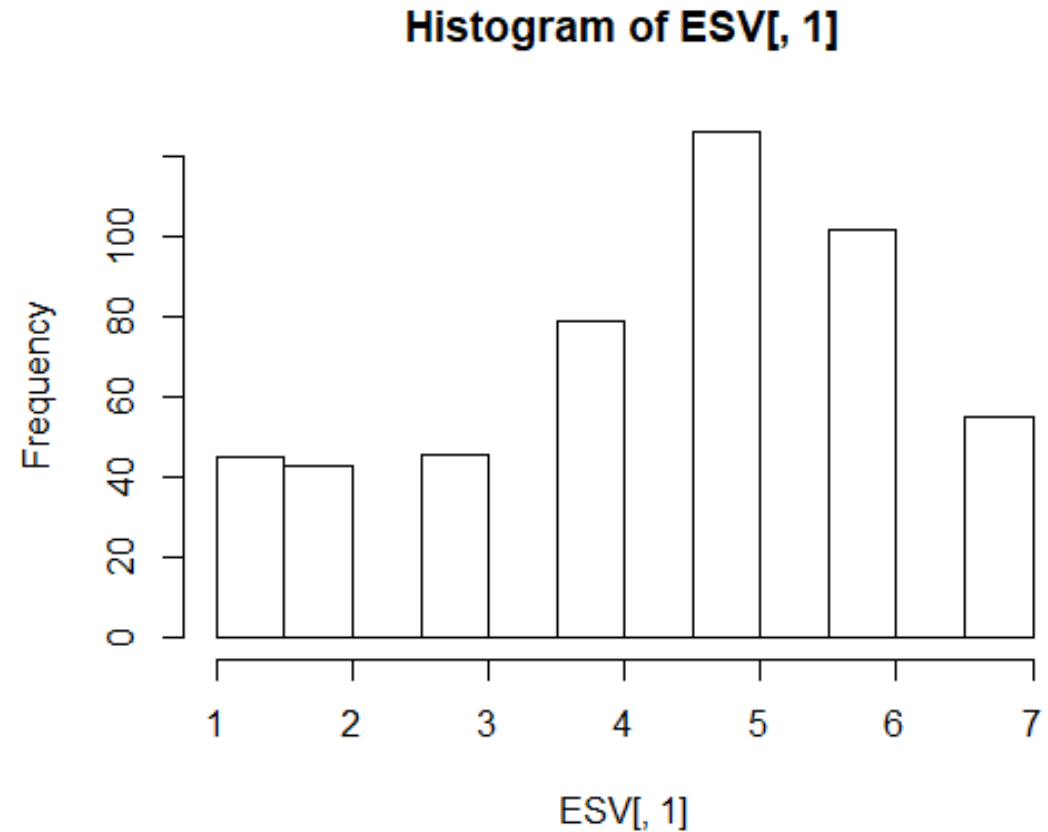
4. ____ Dentro do possível, tenho conseguido as coisas importantes que quero da vida.

5. ____ Se pudesse viver uma segunda vez, não mudaria quase nada na minha vida.

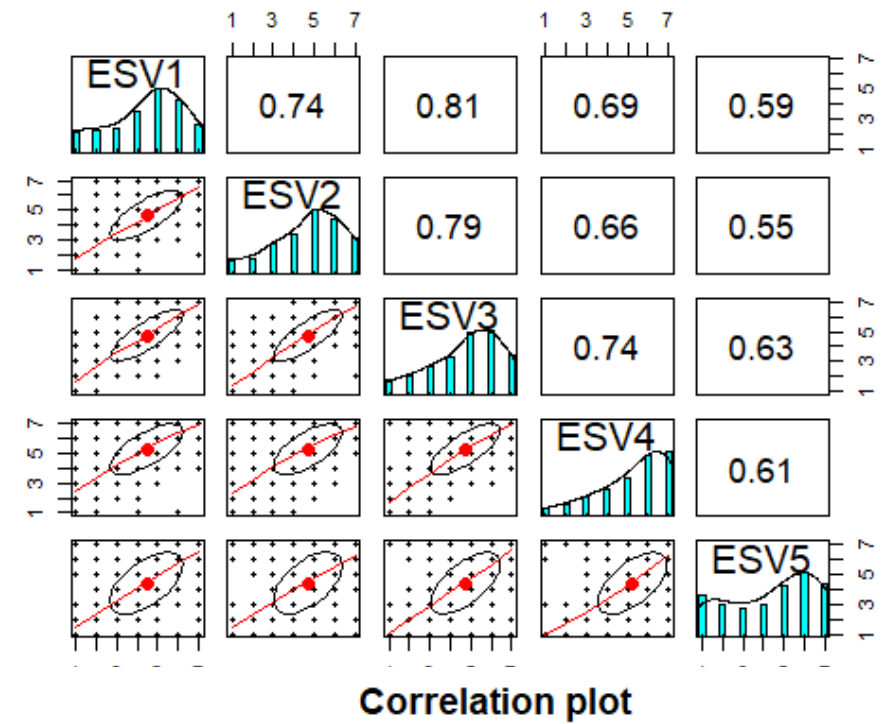
(http://www.vvgouveia.net/sp/images/Gouveia_2005_medindo_a_satisfao_com_a_vida_dos_mdicos_no_brasil.pdf)

Distribuições

- `hist(ESV[,1])`
- `shapiro.test(ESV[,1])`
- Normalidade: Pearson
- Violações da normalidade:
 - Spearman
 - Tetracóricas (dicotômicos)
 - Policóricas (categorias ordenadas)



- `cor.plot(ESV[, -c(6,7)], numbers = TRUE, cex = 0.8)`
- `pairs.panels(ESV[, -c(6,7)])`



Qualidade dos dados

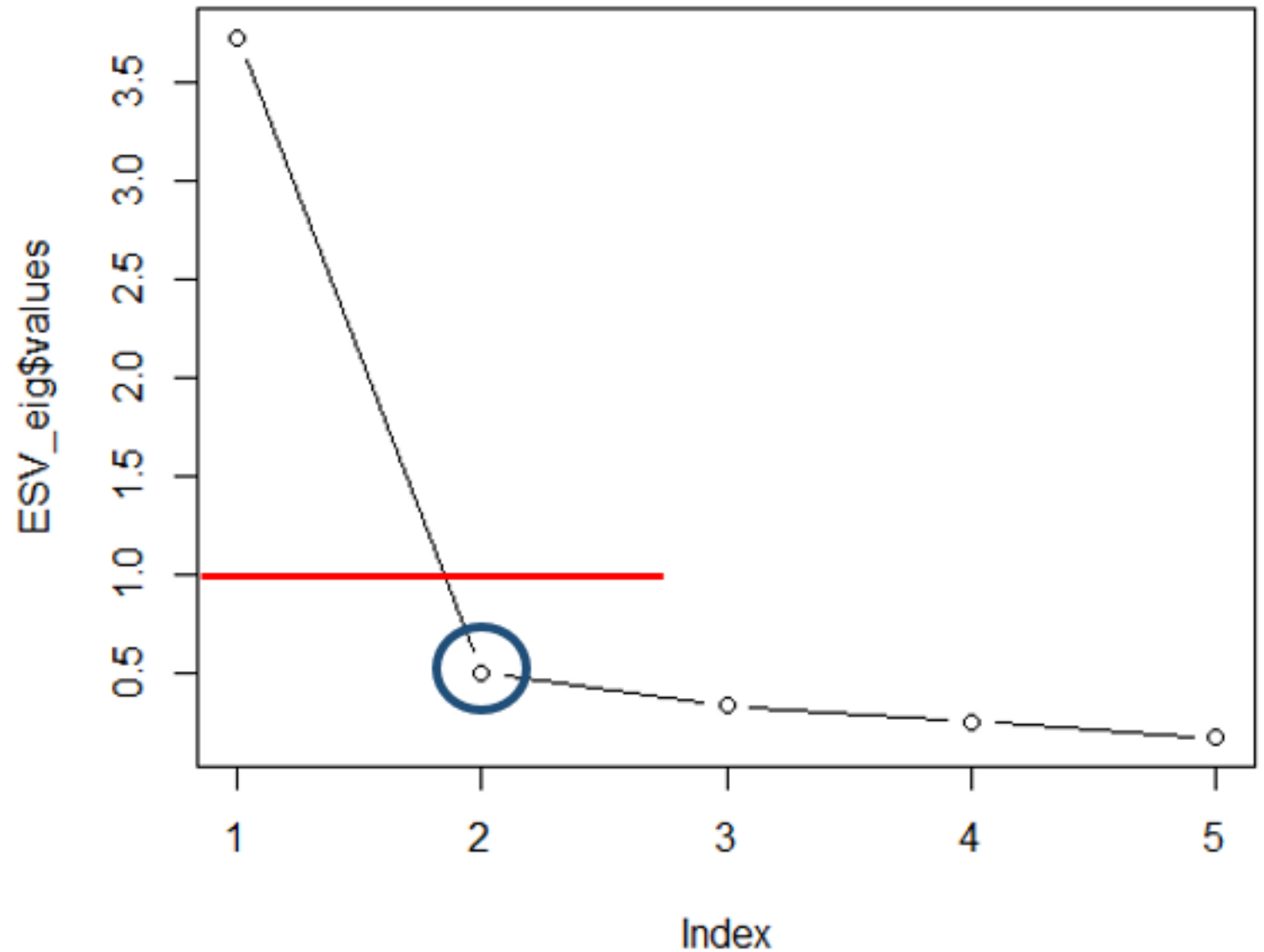
- KMO = proporção entre correlações bivariadas e correlações parciais
 - Quanto mais próximo à unidade, melhor
 - Aceitável acima de 0,6
- Teste de esfericidade de Bartlett ($p < 0,05$)

$$KMO = \frac{\sum_{j \neq k} \sum r_{jk}^2}{\sum_{j \neq k} \sum r_{jk}^2 + \sum_{j \neq k} \sum p_{jk}^2}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Técnicas de retenção de fatores

- Critério de Kaiser (Eigenvalue > 1)
- Critério de Cattell (descontinuidade no gráfico de sedimentação)

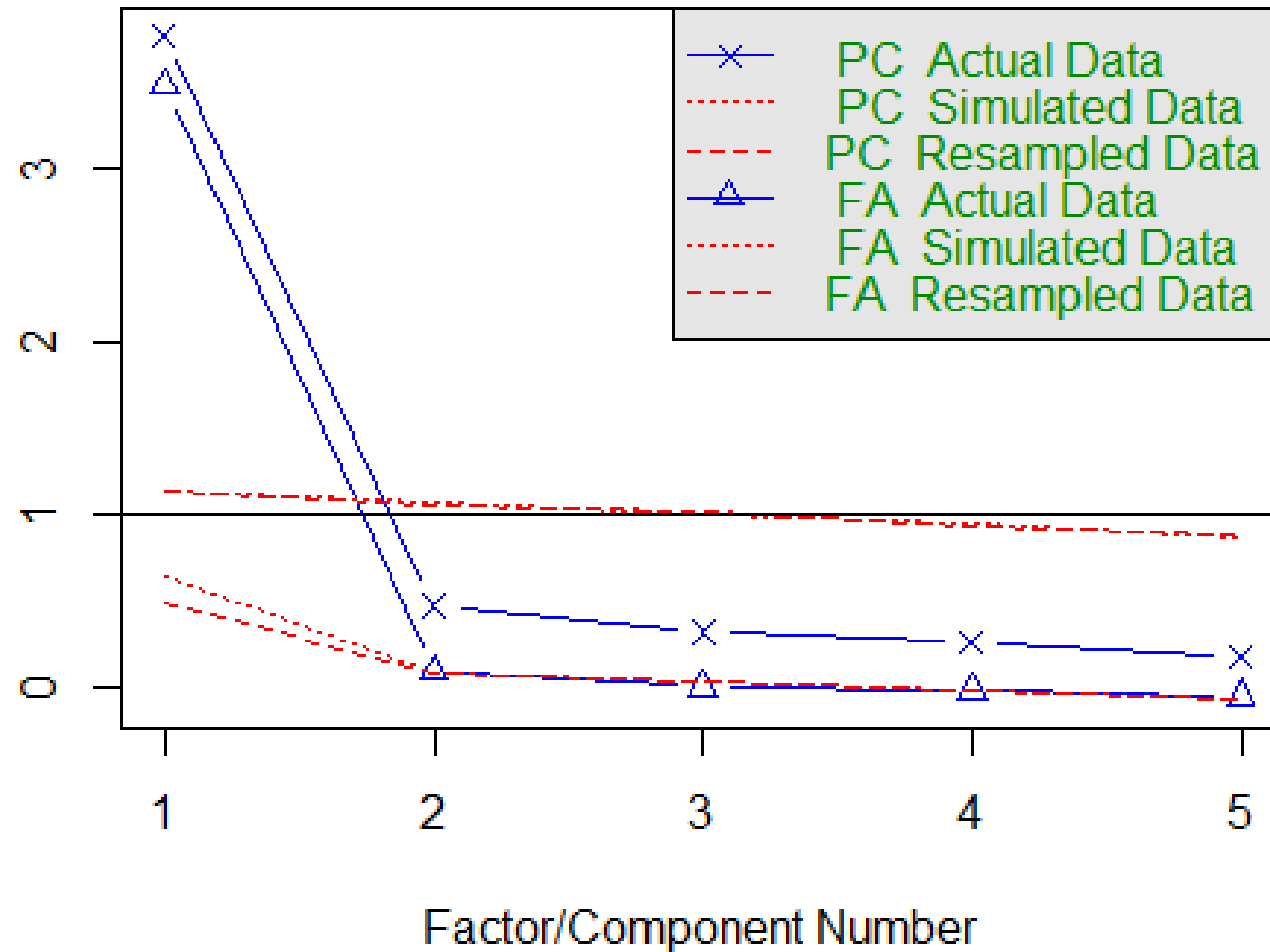


Análise paralela

- Simular dados e comparar o poder explicativo dos fatores
- Simulações: Monte Carlo (paramétrica) e por permutação dos valores

Parallel Analysis Scree Plots

eigenvalues of principal components and factor analysis



```
fa(ESV[,-c(6,7)],cor="poly",fm="minrank")
```

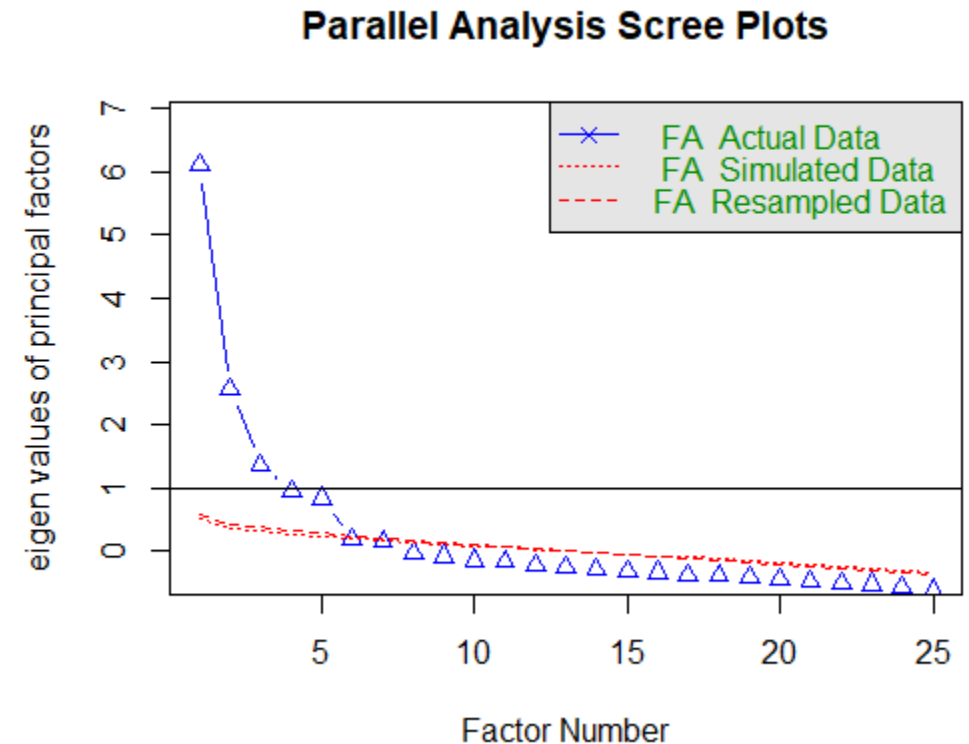
Standardized loadings (pattern matrix) based upon correlation matrix

	MRFA1	h2	u2	com
ESU1	0.87	0.75	0.25	1
ESU2	0.84	0.70	0.30	1
ESU3	0.92	0.85	0.15	1
ESU4	0.82	0.68	0.32	1
ESU5	0.74	0.55	0.45	1

	MRFA1
SS loadings	3.52
Proportion Var	0.70

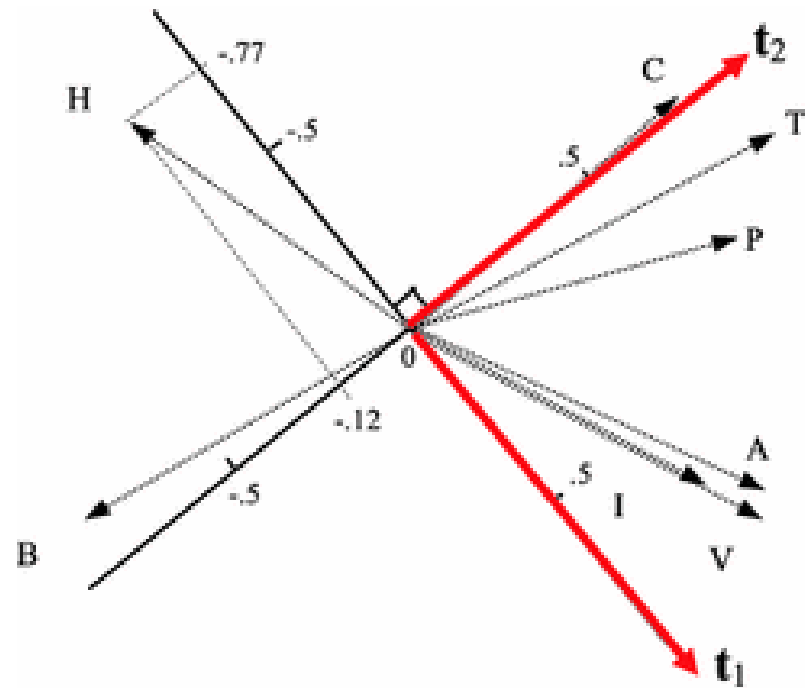
Modelos multidimensionais

- $X = b + a_1F_1 + a_1F_2 + a_2F_1 \dots + e$

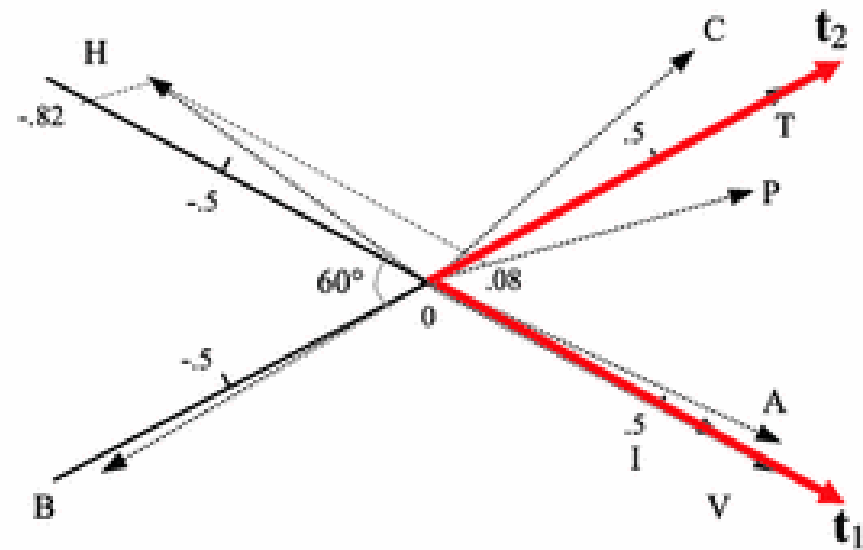


Rotação fatorial

- Rotação fatorial
- A rotação fatorial é uma técnica que visa ajustar a solução fatorial (cargas fatoriais) de modo a deixá-la mais interpretável, ou clara. A solução não rotacionada pode levar aos itens apresentarem correlações com vários fatores. Para diminuir este efeito, os eixos dos fatores são rotacionados no espaço, de modo a permitir (rotação oblíqua) ou não (rotação ortogonal) a correlação entre os fatores.



(a) Orthogonal Rotation



(b) Oblique Rotation

Escores fatoriais

- Fiz a AFE, e agora?
- () somar os itens e usar o resultado em outras análises
- () usa a informação da modelagem estatística, ponderando a importância de cada item, em outras análises

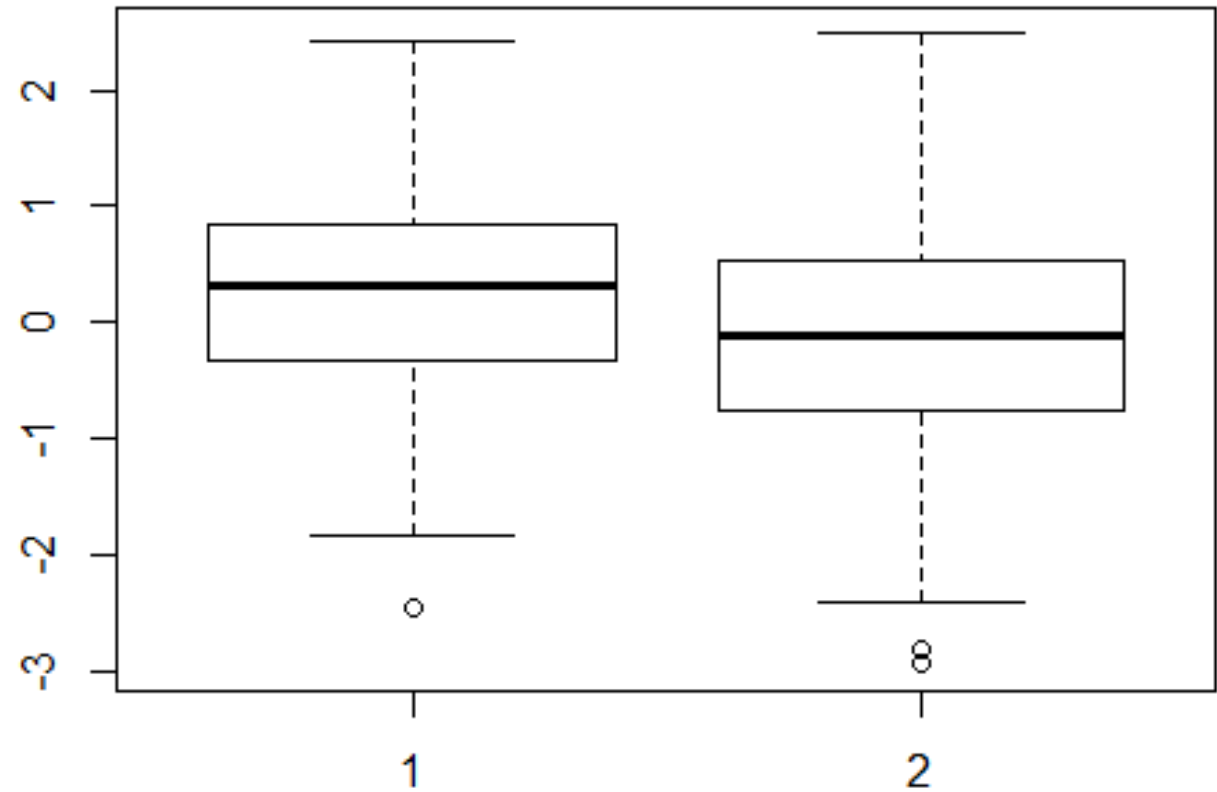
Escores fatoriais

- `Big5_fa<-fa(Big5[,-c(26,27)],5,cor="poly",fm="minrank",rotate = "oblimin",scores = "regression")`
- `View(Big5_fa$scores)`

	MRFA1	MRFA2	MRFA4	MRFA5	MRFA3
1	1.743510658	-0.61406410	-1.38189079	0.75822935	0.68997423
2	NA	NA	NA	NA	NA
3	-0.689141688	-0.00883950	-0.97104036	-1.71044114	-0.09944628
4	-0.500946099	1.07885105	-1.02285449	-0.12124280	1.77448972
5	-1.762444583	0.55404260	-0.53821931	-1.43584878	0.30273618
6	-0.130637187	-0.74167365	-1.62212615	-2.90191335	0.33061680
7	NA	NA	NA	NA	NA
8	0.493125538	-1.04918353	1.25134447	-1.81605333	0.14035380
9	-0.229904915	2.75605764	0.59508649	-0.60651766	-0.17919788

Escores fatoriais

- `t.test(Big5_fa$scores[,5]~Big5$Sexo)`
- `boxplot(Big5_fa$scores[,5]~Big5$Sexo)`





Análise Fatorial Confirmatória e Equações Estruturais: Conceitos Básicos

Prof. Dr. Wagner de Lara Machado

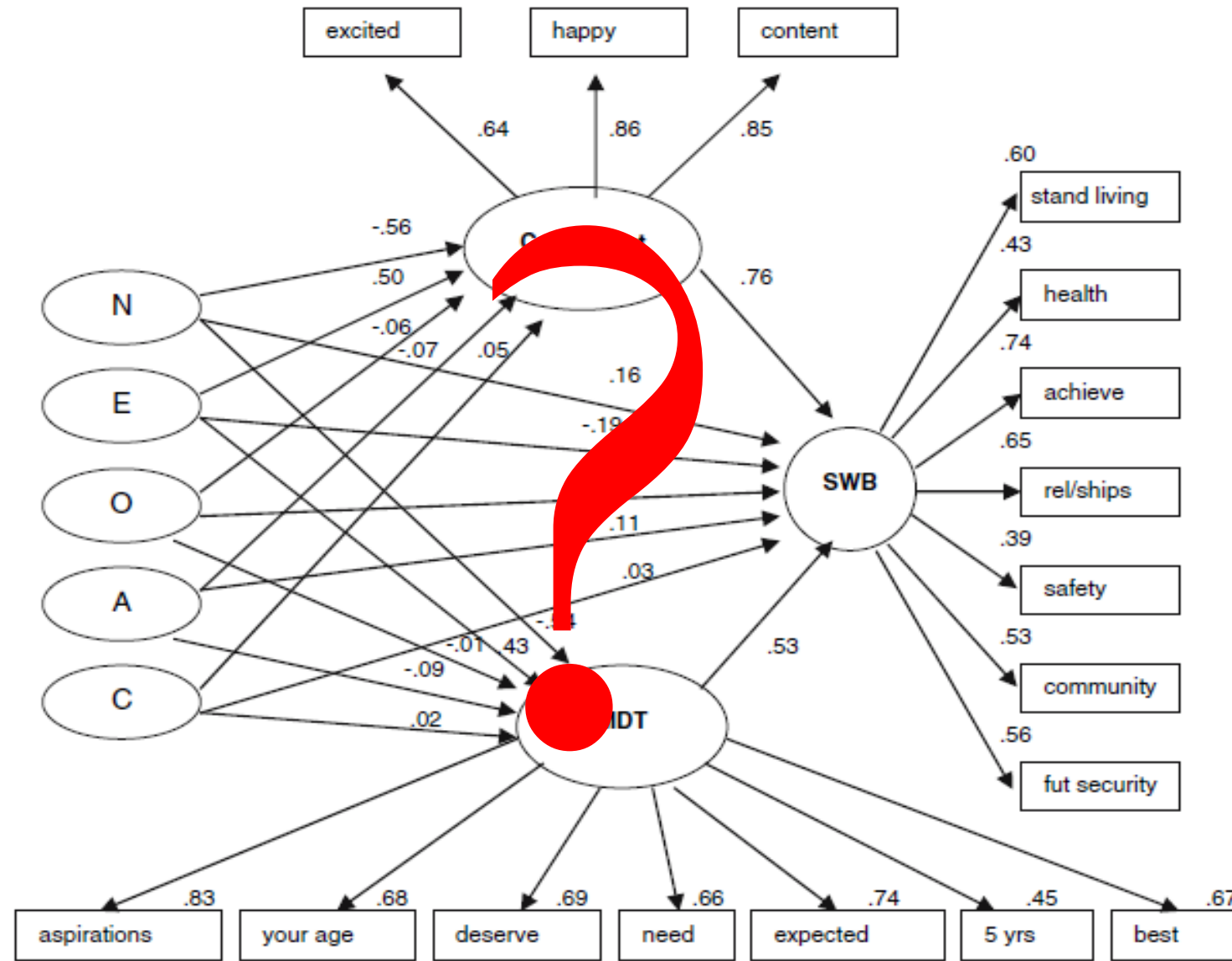


Fig. 3 Personality model of SWB

(Davern, Cummins, & Stokes, 2007)

Análise Fatorial Confirmatória e Equações Estruturais

- Origens
 - Início do século XX
 - Psicometria – Análise Fatorial
 - Genética e Econometria – Análise de regressão múltipla e análise de trajetórias (*path analysis*)
 - Análises multivariadas
 - Necessidade de acomodar uma grande quantidade de variáveis em relações simultâneas
 - Dependência (Regressão múltipla, MANOVA)
 - Interdependência (Análise fatorial exploratória, escalonamento multimensional)

Análise Fatorial Confirmatória e Equações Estruturais

- Características:
 - Relações múltiplas e simultâneas de dependência e interdependência entre variáveis independentes e dependentes
 - Expressas como equações de regressão
 - Expressas como covariâncias (correlações)
 - Definidas pelo pesquisador (modelagem, imposição dos parâmetros)
 - Abordagem confirmatória (teste de modelos)
 - Incorpora a análise de variáveis não diretamente observadas (variáveis latentes, construtos)
 - Estimação dos erros de mensuração
 - Permite a avaliação de efeitos mediadores e moderadores

Análise Fatorial Confirmatória e Equações Estruturais

- Princípio básico
 - Definição de um modelo de relações (regressões, covariâncias, etc)
 - Avaliar o quanto os modelos “se ajustam” (*goodness-of-fit*) aos dados amostrais (matriz de covariâncias ou correlações)
 - Estimar o quanto a matriz amostral representa a matriz populacional (inferência)



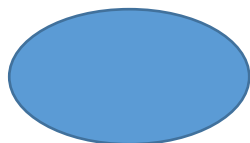
Princípio básico

Análise Fatorial Confirmatória e Equações Estruturais

- Como representar variáveis e relações em um modelo? Diagramas!



Variável observável



Variável latente

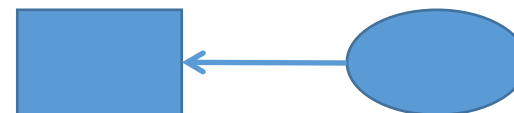


Caminho de regressão

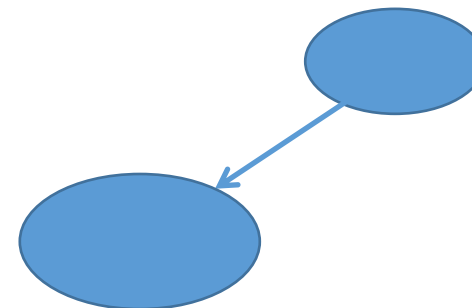


Covariância ou
correlação

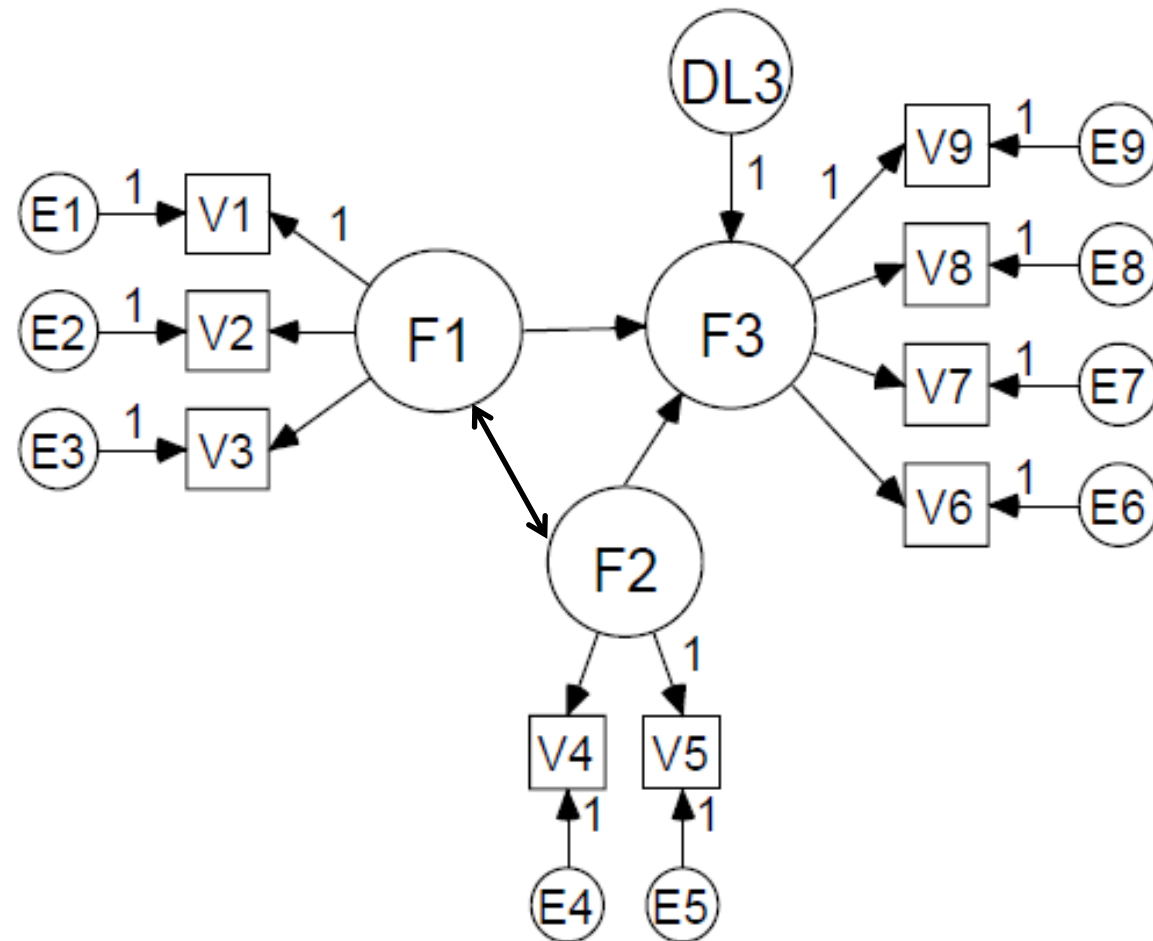
Erro de mensuração



Distúrbio ou resíduo

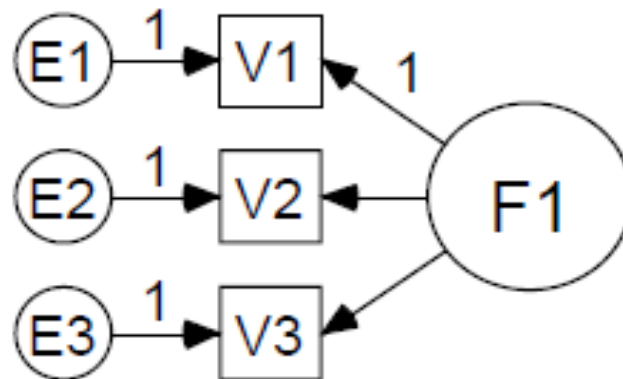


Análise Fatorial Confirmatória e Equações Estruturais



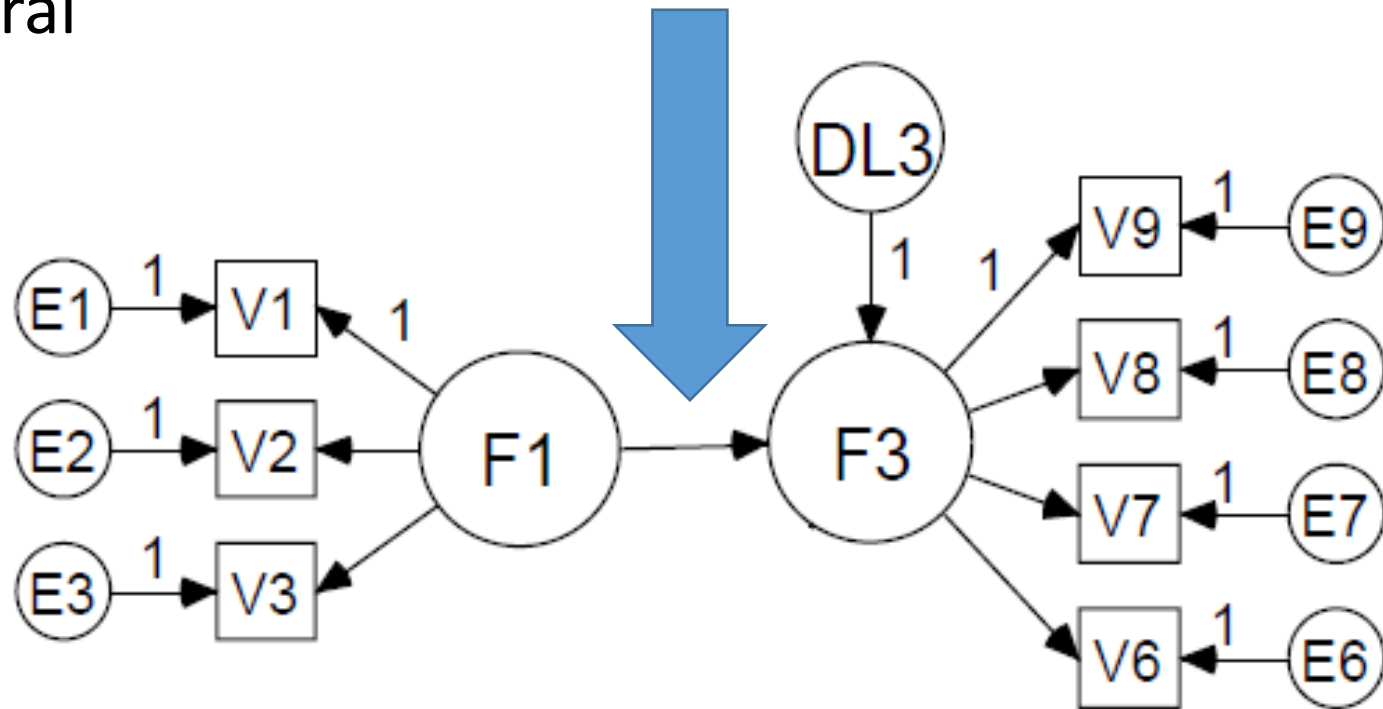
Análise Fatorial Confirmatória e Equações Estruturais

- Níveis das variáveis
 - Exógena – Variáveis independentes (preditor)
 - Endógenas – Variáveis dependentes (predita)
 - Exógena/endógena – cumpre ambos os papéis
- Níveis ou porções dos modelos
 - Mensuração/Análise fatorial



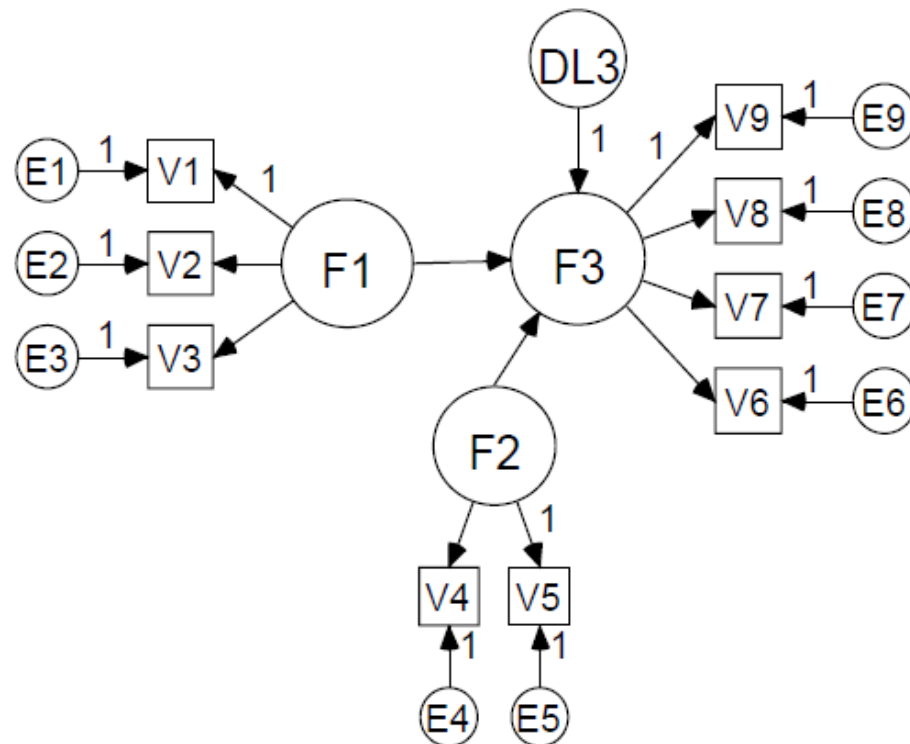
Análise Fatorial Confirmatória e Equações Estruturais

- Níveis ou porções dos modelos
 - Estrutural



Análise Fatorial Confirmatória e Equações Estruturais

- Características dos modelos
 - Modelos Recursivos – relações unidirecionais



Análise Fatorial Confirmatória e Equações Estruturais

- Contextos do uso das Análises de Equações Estruturais
 - Estritamente confirmatório
 - Avaliação de modelos concorrentes (alternativos)*
 - Geração de modelos*****&%@#@#\$\$`&!!!!

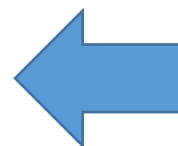
Análise Fatorial Confirmatória e Equações Estruturais

- Métodos de estimação
- Interpretação dos parâmetros
- Índices de ajuste dos modelos
- Processo de decisão
 - Índices de modificação
 - Comparação de modelos concorrentes
- Exercícios práticos

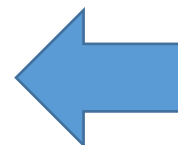


Análise Fatorial Confirmatória e Equações Estruturais

- Métodos de estimação
 - Matrizes de entrada dos dados
 - Variância/Covariância*
 - Correlações de Pearson
 - Correlações Tetracóricas ou Policóricas (dados dicotômicos, ordinais ou violação de normalidade)
 - Métodos de estimação dos modelos (AMOS)
 - Maximum Likelihood*
 - Generalized Least Squares
 - Unweighted Least Squares
 - Scale-free Least Squares
 - Asymptotically Distribution-free



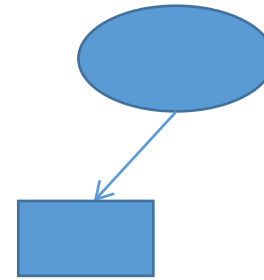
Distribuição normal
multivariada



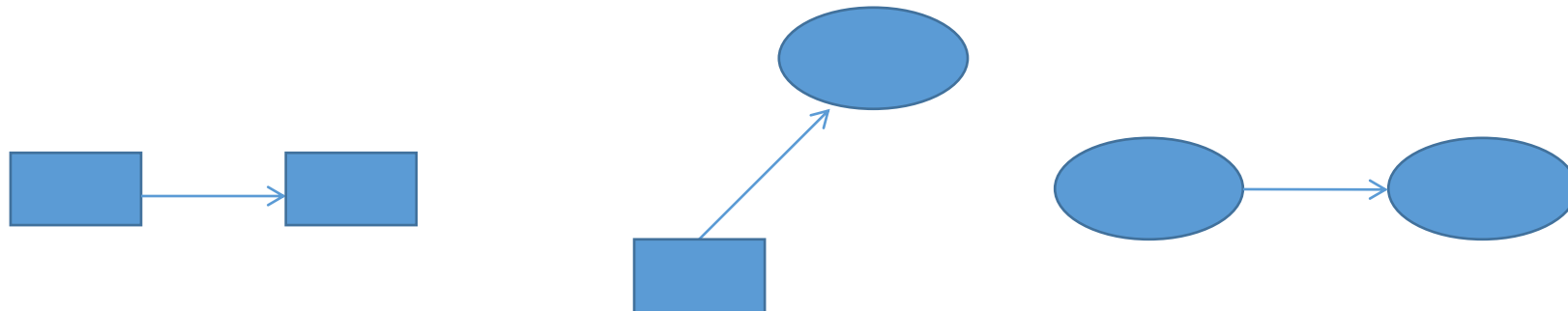
Violações de normalidade


Análise Fatorial Confirmatória e Equações Estruturais

- Interpretação dos parâmetros
 - Como cargas fatoriais, no nível de mensuração



- Como coeficientes angulares (β). Interpretação: quando a VI aumenta 1dp, a VD aumenta/diminui “x” dp.





Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review


JAMES B. SCHREIBER
Duquesne University

AMAURY NORA
University of Houston

FRANCES K. STAGE
New York University

ELIZABETH A. BARLOW
University of Houston

JAMIE KING
Duquesne University



Análise Fatorial Confirmatória e Equações Estruturais

- Índices de ajuste dos modelos

Absolute/predictive fit

Chi-square	χ^2	Ratio of χ^2 to $df \leq 2$ or 3, useful for nested models/model trimming
Akaike information criterion	AIC	Smaller the better; good for model comparison (nonnested), not a single model
Browne–Cudeck criterion	BCC	Smaller the better; good for model comparison, not a single model
Bayes information criterion	BIC	Smaller the better; good for model comparison (nonnested), not a single model
Consistent AIC	CAIC	Smaller the better; good for model comparison (nonnested), not a single model
Expected cross-validation index	ECVI	Smaller the better; good for model comparison (nonnested), not a single model

Análise Fatorial Confirmatória e Equações Estruturais

- Índices de ajuste dos modelos

Comparative fit		Comparison to a baseline (independence) or other model
Normed fit index	NFI	$\geq .95$ for acceptance
Incremental fit index	IFI	$\geq .95$ for acceptance
Tucker–Lewis index	TLI	$\geq .95$ can be $0 > TLI > 1$ for acceptance
Comparative fit index	CFI	$\geq .95$ for acceptance
Relative noncentrality fit index	RNI	$\geq .95$, similar to CFI but can be negative, therefore CFI better choice

Análise Fatorial Confirmatória e Equações Estruturais

- Índices de ajuste dos modelos

Parsimonious fit

Parsimony-adjusted NFI

PNFI

Very sensitive to model size

Parsimony-adjusted CFI

PCFI

Sensitive to model size

Parsimony-adjusted GFI

PGFI

Closer to 1 the better, though typically lower than other indexes and sensitive to model size

Análise Fatorial Confirmatória e Equações Estruturais

- Índices de ajuste dos modelos

Other

Goodness-of-fit index	GFI	$\geq .95$ Not generally recommended
Adjusted GFI	AGFI	$\geq .95$ Performance poor in simulation studies
Hoelter .05 index		Critical N largest sample size for accepting that model is correct
Hoelter .01 index		Hoelter suggestion, $N = 200$, better for satisfactory fit
Root mean square residual	RMR	Smaller, the better; 0 indicates perfect fit
Standardized RMR	SRMR	$\leq .08$
Weighted root mean residual	WRMR	$< .90$
Root mean square error of approximation	RMSEA	$< .06$ to $.08$ with confidence interval

Análise Fatorial Confirmatória e Equações Estruturais

- Referências

- Byrne, B. (2001a). *Structural equation modeling with Amos: Basic concepts, applications and programming*. New Jersey: Lawrence Erlbaum.
- Byrne, B. (2001b). Structural Equation Modeling With AMOS, EQS, and LISREL: Comparative Approaches to Testing for the Factorial Validity of a Measuring Instrument. *International Journal of Testing*, 1 (1), 55-86.
- Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (2005). *Análise multivariada de dados*. 5a. edição. Porto Alegre: Artmed.
- Jackson, D. L., Gillasp Jr., J. A., Purc-Stephenson, R. (2009). Reporting Practices in Confirmatory Factor Analysis: An Overview and Some Recommendations. *Psychological Methods*, 14(1), 6-23.
- Pilati, R. & Laros, J.A. (2007). Modelos de equações estruturais em psicologia: conceitos e aplicações. *Psicologia: Teoria e Pesquisa*, 23(2), 205-216.
- Schreiber, J.B., Stage, F.K., King, J., Nora, A., & Barlow, E.A. (2006). Reporting structural equation modeling and confirmatory factor analysis results: a review. *The Journal of Educational Research*, 99(6), 324-337.

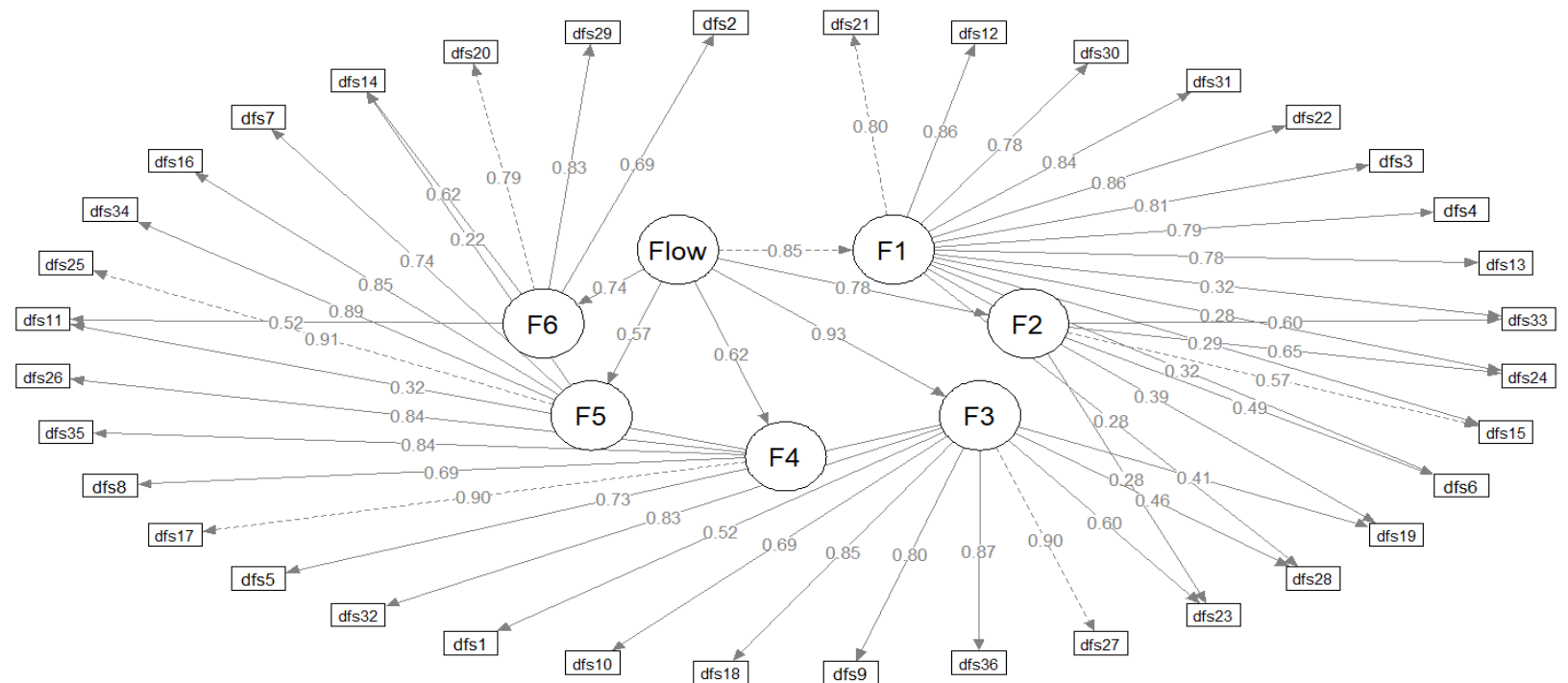
Análise Fatorial Confirmatória e Equações Estruturais

- Aplicações
 - Análise de trajetórias (*path analysis*)
 - Análise fatorial confirmatória*
 - Modelagem latente (*full latent models*)*
 - Análise multigrupos*
 - Análise de curva latente (*latent growth curve*)*
 - Modelagem multinível
 - Matriz multitraço-multimétodo
 - Modelagem de equações estruturais exploratória
 - ...

Análise Fatorial Confirmatória e Equações Estruturais

- Outros pressupostos
 - Tamanho amostral
 - >200 observações, sendo recomendada a razão de 10/1 em relação aos parâmetros estimados
 - Dados omissos (*missing values*) e *outliers*
 - Aumentam o número de parâmetros estimados (médias, interceptos)
 - Aconselhado utilizar algum método de imputação (para ambos)
- Normalidade, homocedasticidade e multicolinearidade
 - Normalidade univariada (Kolmogorov-Smirnov)
 - Normalidade multivariada (Mardia)
 - Homocedasticidade (Levene e *M* de Box)
 - Matriz de correlações

Análise Fatorial Confirmatória em Psicologia



Semelhanças e diferenças AFE & AFC

- Ambos visam identificar fatores que explicam a variação e co-variação entre um conjunto de indicadores
- AFE – Data driven - geralmente um procedimento descritivo ou exploratório
- AFC – Theory Driven - pesquisador deve pré-especificar todos os aspectos do modelo latente: o número de fatores, o padrão de cargas entre indicador e fator, e assim por diante

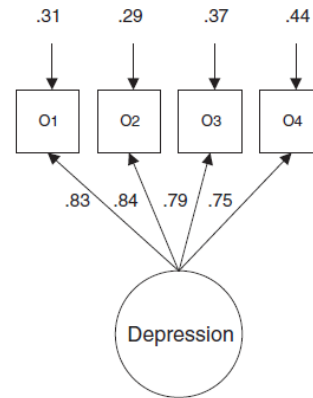


FIGURE 2.1. Path diagram of the one-factor model.

Pacote LAVAAN

- Latent Variable Analysis
- Permite especificação do modelo – O pesquisador indica qual item ou variável devem carregar em cada fator
- Permite fazer análises:
 - multi-grupo
 - investigação de índices de modificação
 - Modelagem por equações – regressão de fatores e mediação



Resultado ESV

User model versus baseline model:

Comparative Fit Index (CFI)	0.995
Tucker-Lewis Index (TLI)	0.990

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-4054.623
Loglikelihood unrestricted model (H1)	-4047.836

Number of free parameters	10
Akaike (AIC)	8129.245
Bayesian (BIC)	8171.311
Sample-size adjusted Bayesian (BIC)	8139.571

Root Mean Square Error of Approximation:

RMSEA	0.059
90 Percent Confidence Interval	0.022 0.097
P-value RMSEA <= 0.05	0.298

Standardized Root Mean Square Residual:

SRMR	0.018
------	-------

Índices de Ajuste do Modelo

Valores de Referência

AIC (*Akaike Information Criterion*), o BIC (*Bayesian Information Criterion*) = Quanto menor melhor!

CFI (*Comparative Fit Index*) & TLI (*Tucker-Lewis Index*) = acima de 0,95 sugerem ajuste excelente, já valores acima de 0,90 indicam que a qualidade de ajuste é satisfatória

RMSEA (*Root Mean Square Error of Approximation*) e SRMR (*Standardized Root Mean Square Residual*) = menores que 0,05 indicam bom ajuste, enquanto valores menores que 0,08 indicam ajuste aceitável

(Hu & Bentler, 1999).



Resultado ESV

Cargas fatoriais padronizadas e
não padronizadas
Significância da carga no fator
Comunalidade

Latent Variables:

SV =~

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
ESV1	1.000				1.535	0.869
ESV2	0.916	0.037	24.778	0.000	1.406	0.845
ESV3	1.029	0.035	29.316	0.000	1.579	0.928
ESV4	0.875	0.039	22.449	0.000	1.344	0.798
ESV5	0.911	0.052	17.611	0.000	1.398	0.683

Carga não-
padronizada

Significância da
carga fatorial

Carga padronizada

R-Square:

	Estimate
ESV1	0.756
ESV2	0.713
ESV3	0.860
ESV4	0.637
ESV5	0.466

Variância explicada
do item pelo
modelo

