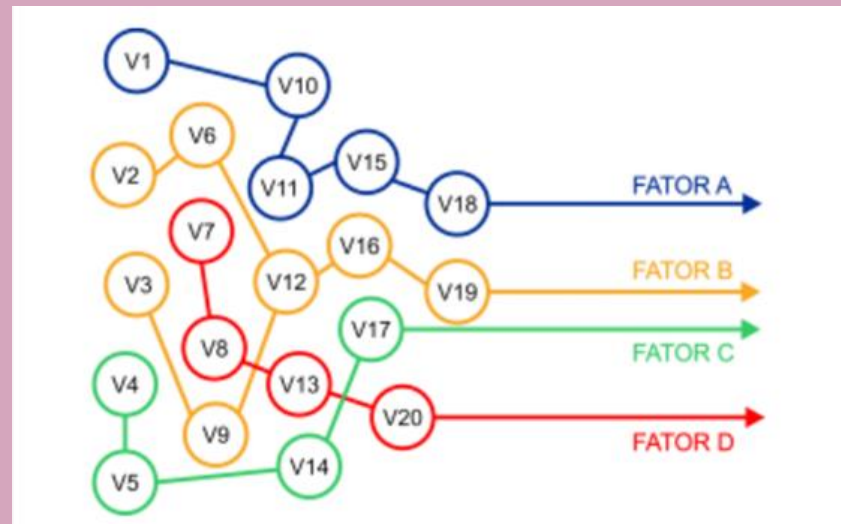


ANÁLISE FATORIAL EXPLORATÓRIA



Tatiele Oliveira
Departamento de Psicologia Social e do Trabalho

ANÁLISE FATORIAL EXPLORATÓRIA

- O que é:
 - A análise fatorial exploratória (AFE) é um conjunto de técnicas estatísticas amplamente utilizadas em pesquisas que utilizem questionários, escalas, inventários, instrumentos, etc.
 - Tem por objetivo encontrar a estrutura 'fatorial' dessas ferramentas, a fim de possibilitar o pesquisador gerar escores a partir dos itens.
- Análise Fatorial, de maneira mais ampla, pode ser dividida em Exploratória (AFE) e Confirmatória (AFC);
- Focaremos especificamente na AFE.

ANÁLISE FATORIAL EXPLORATÓRIA

- **Cuidados com a AFE**

Durante sua execução, diversas decisões precisam ser tomadas a fim de se obter uma estrutura fatorial adequada:

- Diferenças entre análise fatorial exploratória e análise de componentes principais;
- Retenção e rotação fatorial;
- Carga Fatorial;
- Variância explicada;
- Tamanho da amostra ideal

ANÁLISE FATORIAL EXPLORATÓRIA

- O que você compreender desde o início:
 - A análise fatorial exploratória (AFE) não para de evoluir;
 - O SPSS não tem boas opções;
 - Um dos melhores software atualmente ☐ FACTOR;
 - Usaremos o JAMOVl.

ANÁLISE FATORIAL EXPLORATÓRIA

- Objetivo da AFE
 - Conjunto de técnicas multivariadas que tem como objetivo encontrar a estrutura subjacente em uma matriz de dados e determinar o número e a natureza das variáveis latentes (fatores) que melhor representam um conjunto de variáveis observadas (Brown, 2006).
 - Uma variável latente é uma variável não observada que explica o padrão de resposta aos itens.

ANÁLISE FATORIAL EXPLORATÓRIA

- Objetivo da AFE

Eu me sinto exausto só de pensar em trabalhar	1	2	3	4	5
Não tenho tido paciência para lidar com as pessoas no meu trabalho	1	2	3	4	5
Sinto que meu trabalho não tem valor	1	2	3	4	5
Cada dia de trabalho parece uma tortura pra mim	1	2	3	4	5
Tenho me tornado frio com as pessoas as quais preciso lidar no meu trabalho	1	2	3	4	5
Meu trabalho me realiza	1	2	3	4	5
Após um dia de trabalho, não tenho ânimo para mais nada	1	2	3	4	5
Por vezes, chego a sentir raiva das pessoas do meu trabalho	1	2	3	4	5
Meu trabalho complementa positivamente a minha vida	1	2	3	4	5

ANÁLISE FATORIAL EXPLORATÓRIA

- Objetivo da AFE
 - As variáveis observadas “pertencem” a um mesmo fator quando, e se, elas partilham uma variância em comum (i.e., são influenciadas pelo mesmo construto subjacente) (Brown, 2006).
 - Fator é uma variável latente (e.g., autoestima) que influencia mais de uma variável observada
 - ‘Estou satisfeito comigo’
 - ‘Tenho boas qualidades’
 - ‘Sou uma pessoa de valor’

ANÁLISE FATORIAL EXPLORATÓRIA

- Objetivo da AFE
 - A AFE é geralmente conduzida quando:
 - 1) O pesquisador não possui uma teoria prévia subjacente ou evidências empíricas suficientes que explicitem como os itens de determinado instrumento devem ser agrupados e avaliados;
 - 2) Quando o pesquisador quer confirmar ou refutar a estrutura fatorial de determinado instrumento (Brown, 2006).

ANÁLISE FATORIAL EXPLORATÓRIA

- **Análise Fatorial e Evidências de Validade**

- Evidências de Validade Baseada na Estrutura Interna (AERA, APA, & NCME, 2014);
- Pode ser realizada em escalas do tipo Likert e escalas dicotômicas;
 - Pode ser implementada em dados mistos (politômicos e dicotômicos – desde que sejam ordinais);
 - Não deve ser utilizada em itens nominais.

ANÁLISE FATORIAL EXPLORATÓRIA

- **Análises Fatoriais Exploratórias vs Análise de Componentes Principais**

- A AFE e a ACP são duas técnicas que têm por objetivo reduzir um determinado número de itens a um menor número de variáveis;
- A ACP foi por muito tempo o método de redução de dados mais utilizado nas pesquisas em Psicologia;
- Cálculos computacionais da ACP mais simples, mais rápidos e mais baratos;
- ACP é, ainda hoje, o método padrão de redução de dados em muitos dos principais programas estatísticos (por exemplo, SPSS e SAS).

ANÁLISE FATORIAL EXPLORATÓRIA

SPSS Menu: **Analisar** > **Redução de dimensão** > **Fator...**

Análise de fator dialog box:

- Variáveis: (Empty)
- Method: **Componentes principais**
- Analisar: ☒ Matriz de correlação, ☐ Matriz de covariância
- Exibir: ☒ Solução de fator não rotacionado, ☐ Diagrama de sedimentação
- Extrair: ☒ Com base no valor próprio, ☐ Número fixo de fatores
- Valores próprios superiores a: **1**
- Fatores a extrair: (Empty)
- Máximo de iterações por convergência: **25**

Análise de fator: Extração dialog box:

- Método: **Componentes principais**
- Analisar: ☒ Matriz de correlação, ☐ Matriz de covariância
- Exibir: ☒ Solução de fator não rotacionado, ☐ Diagrama de sedimentação
- Extrair: ☒ Com base no valor próprio, ☐ Número fixo de fatores
- Valores próprios superiores a: **1**
- Fatores a extrair: (Empty)
- Máximo de iterações por convergência: **25**

Data Table:

	DGI11	DGI12	DGI13	DGI14	DGI15	DGI16	DGI17	DGI18	DGI19	DGI20	DGI21	DGI22	DGI23
1	4.00	3.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
2	2.00	1.00	4.00	2.00	2.00	5.00	2.00	2.00	5.00	1.00	5.00	2.00	1.00
3	5.00	3.00	5.00	2.00	3.00	5.00	2.00	2.00	5.00	1.00	5.00	4.00	2.00
4	2.00	2.00	5.00	2.00	2.00	5.00	2.00	2.00	5.00	1.00	5.00	4.00	2.00
5	2.00	3.00	5.00	2.00	3.00	5.00	2.00	2.00	5.00	1.00	5.00	4.00	2.00
6	4.00	3.00	4.00	5.00	2.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
7	5.00	2.00	4.00	5.00	2.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
8	5.00	1.00	4.00	5.00	2.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
9	2.00	1.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
10	1.00	4.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
11	3.00	3.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
12	4.00	5.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
13	1.00	1.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
14	2.00	4.00	4.00	5.00	2.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
15	1.00	3.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
16	2.00	2.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
17	2.00	3.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
18	2.00	2.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
19	2.00	2.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
20	2.00	2.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
21	2.00	2.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
22	2.00	2.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
23	2.00	2.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
24	2.00	2.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00
25	2.00	2.00	5.00	2.00	1.00	4.00	5.00	3.00	5.00	1.00	5.00	4.00	2.00

ANÁLISE FATORIAL EXPLORATÓRIA

- **Análises Fatoriais Exploratórias vs Análise de Componentes Principais**
 - AFEs geram fatores;
 - ACPs geram componentes;
 - A diferença entre componentes e fatores está relacionada à forma como os itens são retidos.
 - Ambos os métodos de redução de dados assumem que a variância de uma variável é composta por três aspectos: variância específica; variância comum e variância de erro.

ANÁLISE FATORIAL EXPLORATÓRIA

- **Análises Fatoriais Exploratórias vs Análise de Componentes Principais**

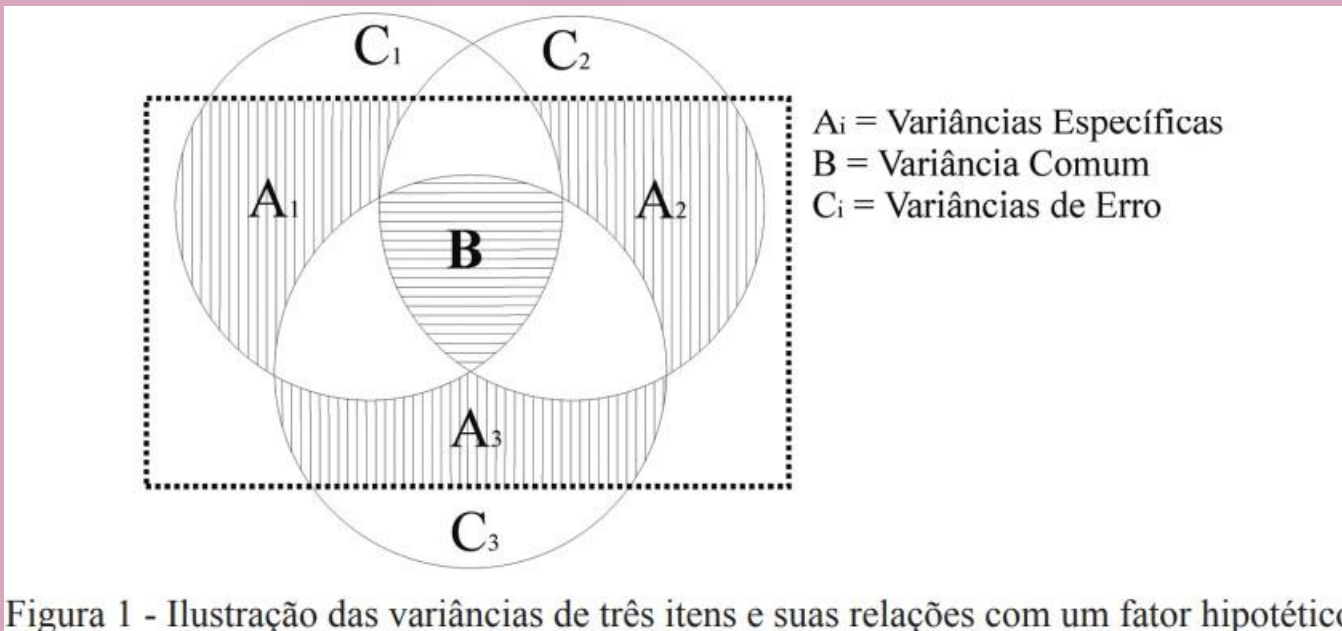


Figura 1 - Ilustração das variâncias de três itens e suas relações com um fator hipotético.

Retirado de Damásio
(2012).

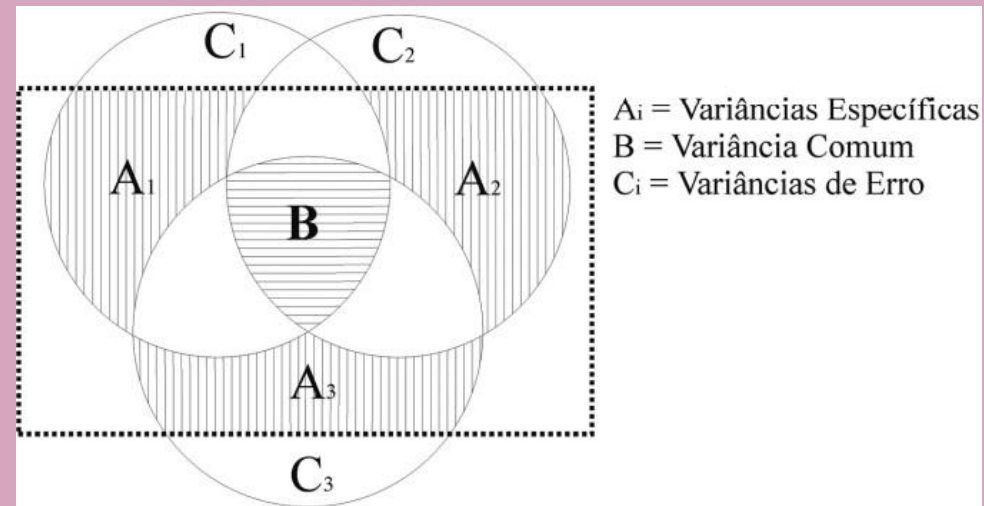
Variância específica: porção de variância do item que não é compartilhada com nenhuma outra variável.

Variância comum: variância que é compartilhada entre todos os itens que compõem determinado fator ou componente.

Variância de erro: parcela do item não explicada pelo componente ou fator.

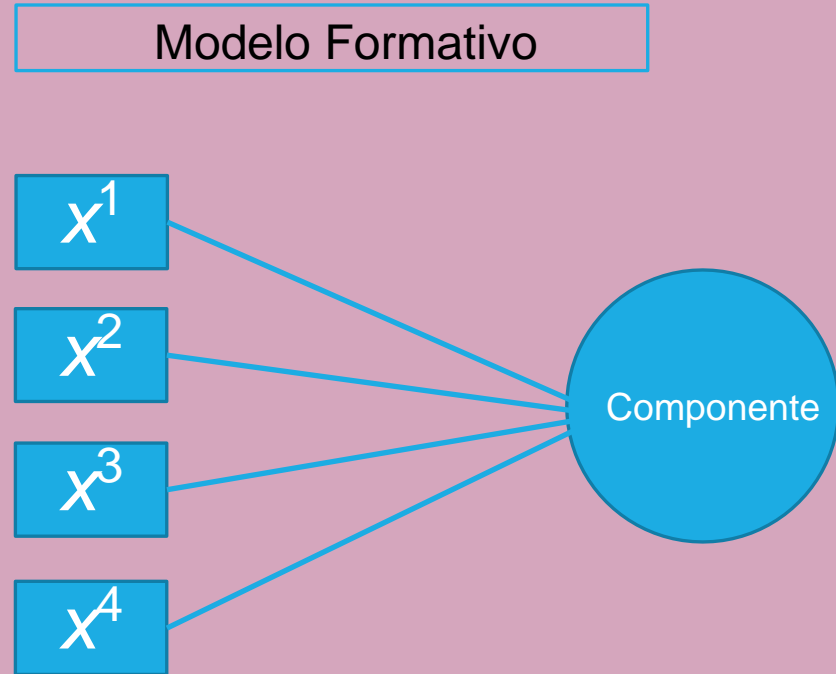
ANÁLISE FATORIAL EXPLORATÓRIA

- **Análises Fatoriais Exploratórias vs Análise de Componentes Principais**
- A ACP está baseada apenas na correlação linear das variáveis observadas, e não diferencia a variância comum da variância específica entre os itens, incluindo ambos os tipos de variância (comum e específica (na Figura 1, $A_i + B$) na análise dos dados.



ANÁLISE FATORIAL EXPLORATÓRIA

■ Análises Fatoriais Exploratórias vs Análise de Componentes Principais



Os indicadores que criam o(s) componente(s)

Exemplos:

Nível socioeconômico

- Escolaridade
- Renda
- Bens materiais
- Emprego (Classificação ISEI)

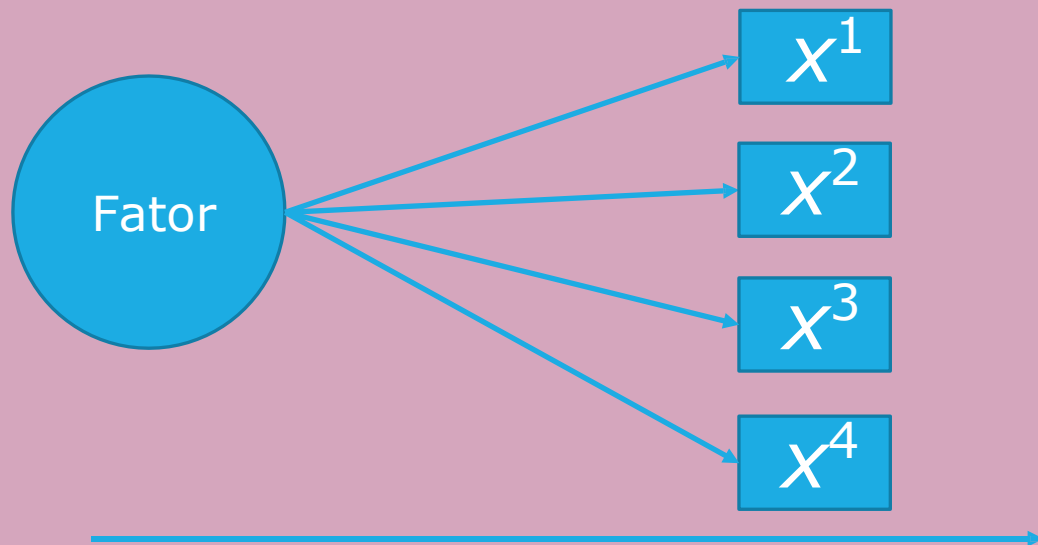
Qualidade de um supermercado

- Preços
- Atendimento dos funcionários
- Espaço físico
- Estacionamento
- Segurança da região
- Variedade de produtos (...)

ANÁLISE FATORIAL EXPLORATÓRIA

■ Análises Fatoriais Exploratórias vs Análise de Componentes Principais

Modelo Reflexivo



O traço latente que gera o comportamento

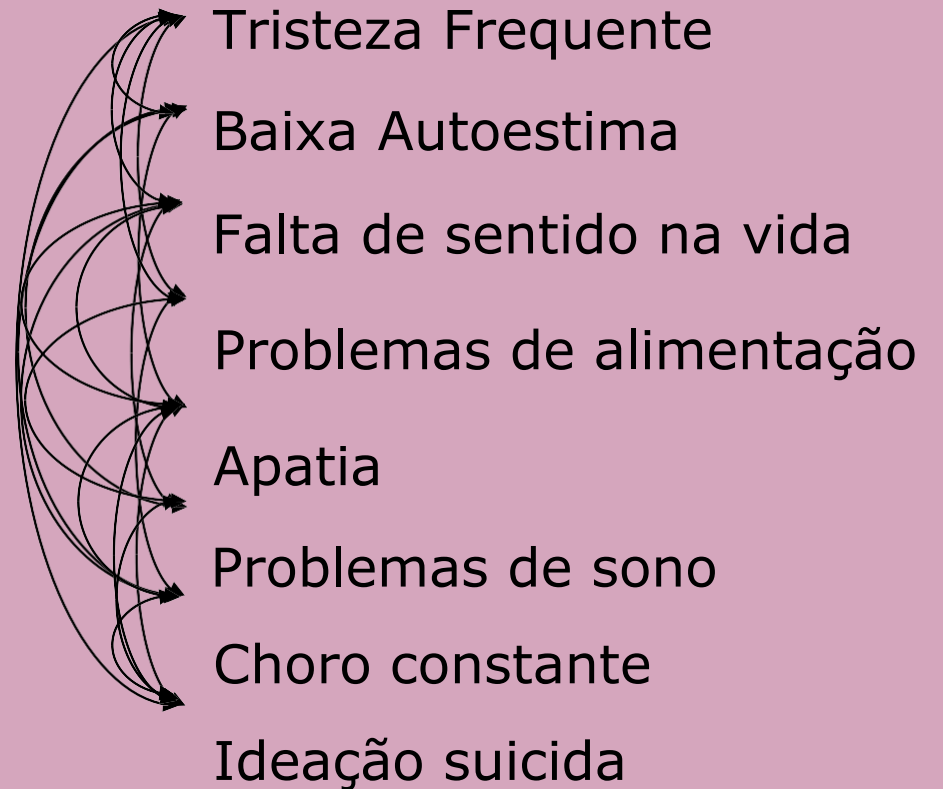
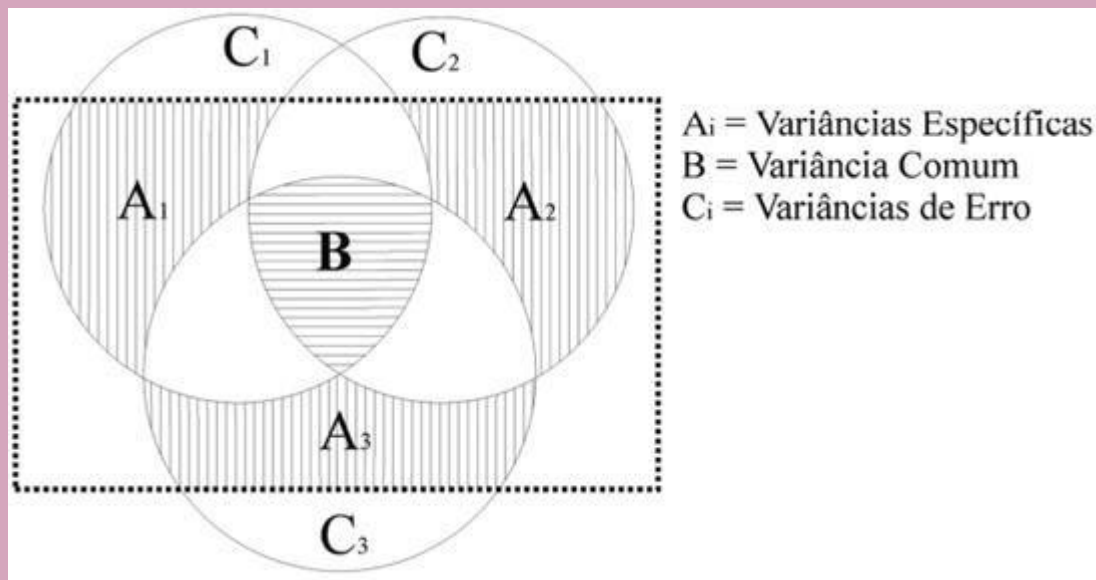
Depressão

- Tristeza Frequente
- Falta de sentido na vida
- Choro constante
- Apatia
- Problemas de sono
- Problemas de alimentação
- Baixa Autoestima
- Ideação suicida

ANÁLISE FATORIAL EXPLORATÓRIA

■ Análises Fatoriais Exploratórias vs Análise de Componentes Principais

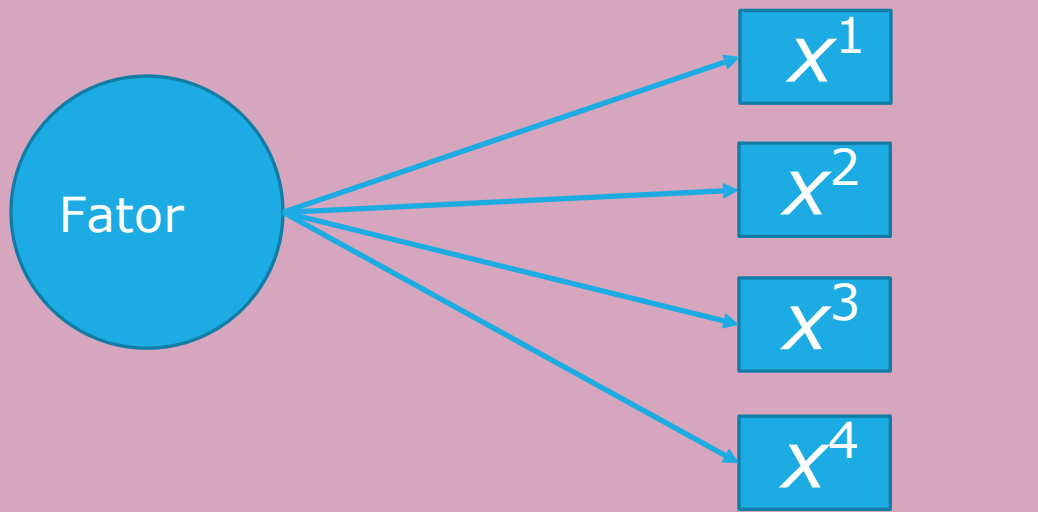
Exemplos: Depressão



ANÁLISE FATORIAL EXPLORATÓRIA

▪ Análises Fatoriais Exploratórias vs Análise de Componentes Principais

Modelo Reflexivo



O traço latente que gera o comportamento

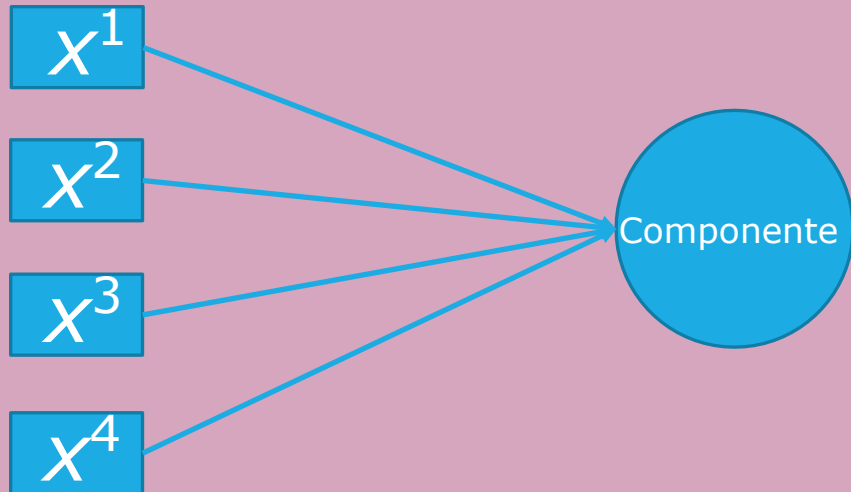
Depressão

- Tristeza Frequente
- Falta de sentido na vida
- Choro constante
- Apatia
- Problemas de sono
- Problemas de alimentação
- Baixa Autoestima
- Ideação suicida

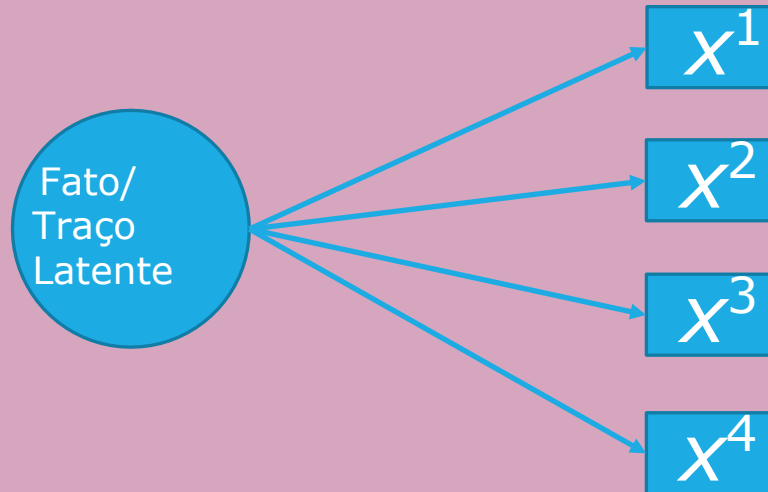
ANÁLISE FATORIAL EXPLORATÓRIA

- Análises Fatoriais Exploratórias vs Análise de Componentes Principais**

Componentes Principais
Modelo Formativo



Análise Fatorial
Modelo Reflexivo



ANÁLISE FATORIAL EXPLORATÓRIA

- **Análises Fatoriais Exploratórias vs Análise de Componentes Principais**
- Em termos práticos...
 - São duas técnicas com racionais teóricos subjacentes completamente distintos, embora seja similar apenas em seu objetivo: reduzir um número maior de indicadores a um número menor de variáveis (componentes ou fatores);
 - A ACP tende a gerar maior carga componencial dos itens, quando comparado com a AF, porque mais fontes de variância são adicionadas;
 - É muito difícil conceber modelos formativos (componentes) na Psicologia. Mas em outras áreas (administração, economia, etc.) é totalmente plausível.



FATORABILIDADE DA MATRIZ

ANÁLISE FATORIAL EXPLORATÓRIA

- **Kaiser Meyer-Olkmin (KMO)**

- Índice de adequação da amostra;
- Avalia se a matriz de dados é passível de fatoração;
- Sugere a proporção de variância dos itens que pode estar sendo explicada por uma variável latente.

(Lorenzo-Seva, Timmerman & Kiers, 2011).

- Valores variam de 0 a 1.
 - $< 0,5$ são considerados inaceitáveis;
 - $> 0,5$ e $\leq 0,7$ são considerados medíocres;
 - $> 0,7$ e $\leq 0,8$ são considerados bons;
 - $> 0,8$ e $0,9$ são considerados ótimos e excelentes.

(Hutcheson & Sofroniou, 1999)

ANÁLISE FATORIAL EXPLORATÓRIA

- **Teste de esfericidade de Barlett**

- Avalia em que medida a matriz de (co)variância é similar a uma matriz-identidade (i.e., os elementos da diagonal principal tem valor igual a um, e os demais elementos da matriz são aproximadamente zero, ou seja, não apresentam correlações entre si, [Field, 2020]).
- Valores do teste de esfericidade de Bartlett com níveis de significância $p < 0,05$ indicam que a matriz é fatorável (Tabachnick & Fidell, 2007);

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

RETENÇÃO FATORIAL

ANÁLISE FATORIAL EXPLORATÓRIA

▪ **Retenção Fatorial**

- Uma das mais importantes decisões a ser tomada;
- Uma extração inadequada impossibilita a interpretação dos resultados de maneira apropriada;

(Hayton, Allen & Scarpello, 2004).

- 1) Superestimação de fatores
 - Resultados não-parcimoniosos, baseados em construtos supérfluos, com reduzido ou inadequado poder explicativo.

(Patil e cols., 2008)

- 2) Subestimação de fatores
 - Subestimação de fatores retidos resulta em perda significativa de informação.

(Franklin, Gibson, Robertson, Pohlmann & Fralish, 1995)

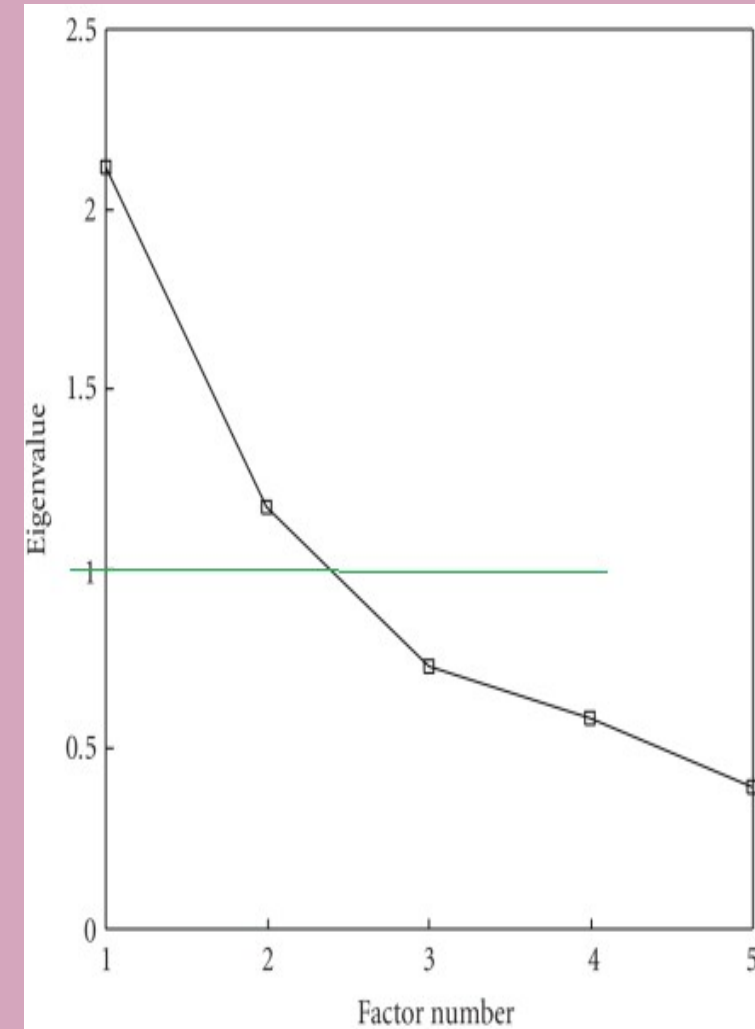
ANÁLISE FATORIAL EXPLORATÓRIA

- **Retenção Fatorial**

- *Eigenvalue* > 1 (o fator explica mais do que 1 item sozinho)
- Ótimos métodos de retenção
 - Análise Paralela
 - Método Hull
 - BIC
 - *Exploratory Graph Analysis*

ANÁLISE FATORIAL EXPLORATÓRIA

- **Eigenvalue > 1**
 - Possivelmente um dos critérios mais utilizados até os dias atuais;
 - Cada fator retido apresenta um eigenvalue que se relaciona ao total de variância explicada por este fator;
 - Fator com eigenvalue < 1 apresenta um total de variância explicada menor do que um único item;
 - Apenas fatores com eigenvalue > 1 são retidos
- (Floyd & Widaman, 1995).



ANÁLISE FATORIAL EXPLORATÓRIA

▪ Eigenvalue > 1

- Apesar da simplicidade, da objetividade e do amplo uso desse critério, há forte consenso na literatura de que seus resultados são imprecisos; (Costello & Osbourne, 2005; Floyd & Widaman; 1995; Patil e cols., 2008; Reise, Waller & Comrey, 2000).
- Critério do eigenvalue > 1 superestimou em 36% dos casos o número de fatores retidos; (Costello & Osbourne, 2005)
- O critério do eigenvalue > 1 tende a superestimar o número de fatores a ser retido devido ao erro amostral; (Lahey, 2010; Ledesma & ValeroMora, 2007).
- O critério de Kaiser-Guttman como método de retenção fatorial não é recomendado. (Patil e cols., 2008; Velicer e cols., 2000).

ANÁLISE FATORIAL EXPLORATÓRIA

▪ Análise Paralela



A AP é um procedimento estatístico de **simulação** Monte-Carlo que compara os eigenvalues dos dados originais, com eigenvalues gerados por matrizes aleatórias construídas no processo de simulação.

A base aleatória possui:

- a) mesma dimensionalidade;
- b) o mesmo número de variáveis;
- c) o mesmo número n de sujeitos do conjunto de dados reais.

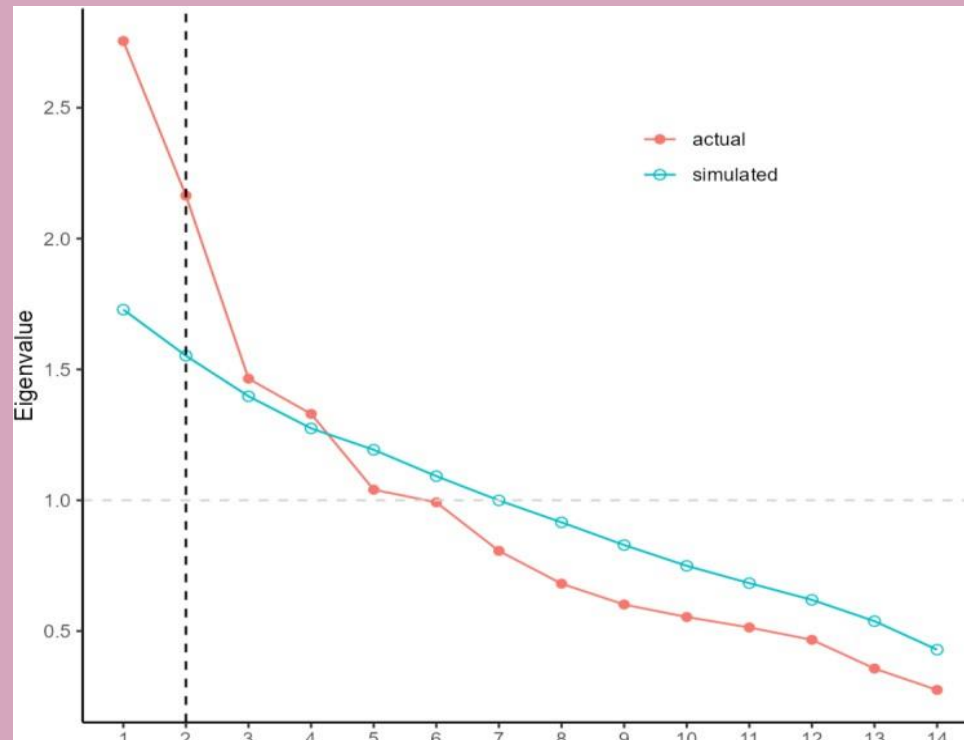
(Laros, 2004)

John Horn
(1928 - 2006)

ANÁLISE FATORIAL EXPLORATÓRIA

■ Análise Paralela

- Os resultados das matrizes aleatórias são contrastados com os dados originais e passa-se a reter apenas os fatores dos dados originais que sejam 'mais fortes' que dos dados aleatórios.



ROTAÇÃO FATORIAL

ANÁLISE FATORIAL EXPLORATÓRIA

- **ROTAÇÃO DE FATORES**

- As rotações fatoriais têm o objetivo de facilitar a interpretação dos fatores, visto que muitas vezes as variáveis analisadas apresentam cargas fatoriais elevadas em mais de um fator.
- As rotações fatoriais podem ser de duas ordens:
 - Ortogonais e Oblíquas

ANÁLISE FATORIAL EXPLORATÓRIA

ROTAÇÃO ORTOGONAL

- Assumem que os fatores extraídos são independentes uns dos outros (i.e, não apresentam correlações entre si; $r = 0$).

Ex: Varimax

- Fatores totalmente independentes uns dos outros;
- Pressuposto raramente obtido nas pesquisas das ciências humanas e da saúde;
- Aspectos humanos e sociais (comportamentos; sintomas; **entre outros**) poucas vezes são divididos em unidades que funcionam independentes umas das outras;
- Métodos ortogonais, em geral, resultam em perda de confiabilidade se os fatores forem correlacionados;
- Tendem à superestimação de variância explicada, visto que possíveis interseções entre os diferentes fatores não são consideradas.

ANÁLISE FATORIAL EXPLORATÓRIA

ROTAÇÃO OBLÍQUA

- Permitem que os fatores se correlacionem entre si;
 - Ex: Oblimin
- Não delimitam a interação entre os fatores *a priori*;
- Se os fatores não forem correlacionados entre si, os resultados obtidos mediante as rotações oblíquas serão bastante semelhantes aos que seriam obtidos por meio das rotações ortogonais.

(Fabrigar e colaboradores, 1999; Sass & Schmitt, 2010).

CARGA FATORIAL

ANÁLISE FATORIAL EXPLORATÓRIA

- **Carga fatorial**

- Medida de importância do item para um determinado fator;
- É baseada na parcela da variância comum do item com os outros itens;
- Carga fatorial (padronizada) varia de $|0 \text{ a } 1|$;
 - Eu tenho objetivos claros na minha vida (+)
 - Sei exatamente o que quero conquistar no meu futuro (+)
 - Minha vida é vazia em termos de metas (-)
- Quanto maior a carga fatorial, mais relevante é aquele item (naquela amostra!);
 - Em geral, se retém itens com cargas fatoriais $< 0,30$.

Tabela 2

Classificação dos fatores encontrados na aplicação do *Self-Reporting Questionnaire* (SRQ-20) na população de trabalhadores de Feira de Santana, Bahia, Brasil, 2001.

Questões do SRQ-20	Trabalhadores			
	Fator I	Fator II	Fator III	Fator IV
Humor depressivo-ansioso				
Sente-se triste ultimamente?	0,7353	-	-	-
Você dorme mal?	0,4709	-	-	-
Você chora mais que de costume?	0,8206	-	-	-
Sente-se nervoso, tenso, preocupado?	0,5303	-	-	-
Tem tremores nas mãos?	0,4369	-	-	-
Assusta-se com facilidade?	0,4506	-	-	-
Decréscimo de energia vital				
O seu trabalho traz sofrimento?	-	0,8484	-	-
Você se cansa com facilidade?	-	0,5970	-	-
Sente-se cansado todo o tempo?	-	0,5953	-	-
Tem dificuldade de ter satisfação em suas tarefas?	-	0,8079	-	-
Sintomas somáticos				
Você sente desconforto estomacal?	-	-	0,8878	-
Você tem falta de apetite?	-	-	0,4992	-
Você tem má digestão?	-	-	0,8396	-
Tem dores de cabeça frequentemente?	-	-	0,4291	-
Pensamentos depressivos				
Tem dificuldade de tomar decisão?	-	-	-	0,6810
Tem perdido interesse pelas coisas?	-	-	-	0,6517
Sente-se inútil em sua vida?	-	-	-	0,6776
Tem dificuldade de pensar claramente?	-	-	-	0,5886
Sente-se incapaz de desempenhar papel útil em sua vida?	-	-	-	0,6919
Tem pensado em dar fim à sua vida?	-	-	-	0,4739

VARIÂNCIA EXPLICADA

ANÁLISE FATORIAL EXPLORATÓRIA

▪ VARIÂNCIA EXPLICADA

- É o quanto o fator explica do padrão de covariância entre os itens
- O que é uma variância explicada alta? O que é baixa?
- Nenhum instrumento será capaz de estimar 100% da variância de um construto hipotético
 - Se não chega a 100%, como se interpreta o número?

***Quanto maior a variância explicada, melhor o seu estudo. Significa que a correlação entre os itens é forte... Logo, o construto está sendo bem mensurado!**

Table 1. Percentage of total variance accounted for by all factors

Decile	Percentage
10th	34
20th	42
30th	49
40th	53
50th	58
60th	61
70th	65
80th	69
90th	76

(Peterson, 2000)

TAMANHO AMOSTRAL

ANÁLISE FATORIAL EXPLORATÓRIA

- Aspecto de divergência de opiniões e controvérsias na literatura científica, principalmente no que diz respeito ao número mínimo de sujeitos necessários;
- Gorsuch (1983) e Hair et al. (2005)
 - $N \geq 100$ sujeitos e um número mínimo de cinco respondentes por item.
- Cattell (1978)
 - $N \geq 250$, 3 a 6 respondentes por item
- Everitt (1975)
 - 10 respostas para cada item avaliado

Não houve experimentos de simulação que corroborassem ou refutassem estas informações.

(MacCallum, Widaman, Zhang, & Hong, 1999)

ANÁLISE FATORIAL EXPLORATÓRIA

- **O que se sabe hoje:**

- Amostras grandes tende a fornecer resultados mais precisos, diminuindo o efeito do erro amostral;
- Fornecem resultados mais próximos ao índice populacional, tanto no que se refere à estrutura fatorial, quanto à carga fatorial e à comunalidade dos itens;

(MacCallum e colaboradores, 1999).

Porém, em relação ao **tamanho mínimo...**

- Barrett e Kline (1981) demonstraram que soluções fatoriais estáveis foram encontradas com um número variando de 1,2 a 3 respondentes por item
- Resultados semelhantes foram encontrados por MacCallum e colaboradores (1999).

- **O que está por trás é mais a qualidade do instrumento do que o N;**

ANÁLISE FATORIAL EXPLORATÓRIA

- **Em casos de construção de instrumento e adaptação, como saber se o instrumento possui boas qualidades psicométricas?**
- **Sobredeterminação:**
 - O grau em que um fator é claramente representado por um número suficiente de itens e pela qualidade (nível) das suas cargas fatoriais e comunalidades.
 - Se um fator é representado por um bom número de itens (tipicamente quatro ou mais);
 - Se estes itens tendem a ser fortemente explicados pelo fator (i.e., apresentam cargas fatoriais elevadas, $> 0,60$);
- Número de respondentes tende a ser menos importante na obtenção de uma boa estrutura fatorial.

ANÁLISE FATORIAL EXPLORATÓRIA

Exploratory Factor Analysis With Small Sample Sizes

J. C. F. de Winter,* D. Dodou,* and P. A. Wieringa

*Department of BioMechanical Engineering, Faculty of Mechanical,
Maritime and Materials Engineering, Delft University of Technology,
The Netherlands*

Exploratory factor analysis (EFA) is generally regarded as a technique for large sample sizes (N), with $N = 50$ as a reasonable absolute minimum. This study offers a comprehensive overview of the conditions in which EFA can yield good quality results for N below 50. Simulations were carried out to estimate the minimum required N for different levels of loadings (λ), number of factors (f), and number of variables (p) and to examine the extent to which a small N solution can sustain the presence of small distortions such as interfactor correlations, model error, secondary loadings, unequal loadings, and unequal p/f . Factor recovery was assessed in terms of pattern congruence coefficients, factor score correlations, Heywood cases, and the gap size between eigenvalues. A subsampling study was also conducted on a psychological dataset of individuals who filled in a Big Five Inventory via the Internet. Results showed that when data are well conditioned (i.e., high λ , low f , high p), EFA can yield reliable results for N well below 50, even in the presence of small distortions. Such conditions may be uncommon but should certainly not be ruled out in behavioral research data.

Applied Psychometrics: Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General

Theodoros A. Kyriazos

Department of Psychology, Panteion University, Athens, Greece
Email: th.kyriazos@gmail.com

How to cite this paper: Kyriazos, T. A. (2018). Applied Psychometrics: Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General. *Psychology*, 9, 2207-2230.
<https://doi.org/10.4236/psych.2018.98126>

Received: July 26, 2018
Accepted: August 21, 2018
Published: August 24, 2018

Copyright © 2018 by author and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Abstract

Adequate statistical power contributes to observing true relationships in a dataset. With a thoughtful power analysis, the adequate but not excessive sample could be detected. Therefore, this paper reviews the issue of what sample size and sample power the researcher should have in the EFA, CFA, and SEM study. Statistical power is the estimation of the sample size that is appropriate for an analysis. In any study, four parameters related to power analysis are Alpha, Beta, statistical power and Effect size. They are prerequisites for a priori sample size determination. Scale development in general and Factor Analysis (EFA, CFA) and SEM are large sample size methods because sample affects precision and replicability of the results. However, the existing literature provides limited and sometimes conflicting guidance on this issue. Generally, for EFA the stronger the data, the smaller the sample can be for an accurate analysis. In CFA and SEM parameter estimates, chi-square tests and goodness of fit indices are equally sensitive to sample size. So the statistical power and precision of CFA/SEM parameter estimates are also influenced by sample size. In this work after reviewing existing sample power analysis rules along with more elaborated methods (like Monte Carlo simulation), we conclude with suggestions for small samples in factor analysis found in literature.

Keywords

Sample Size, Sample Power, SEM, CFA, EFA, Psychometrics, Monte Carlo Simulation, Test Development

REFERÊNCIAS

- Claeskens G, Hjort NL. Model Selection and Model Averaging. Cambridge: University Press; 2008
- Green S. B., Yang Y. (2009a). Commentary on coefficient alpha: a cautionary tale. *Psychometrika* 74, 121–135. 10.1007/s11336-008-9098-4
- McDonald R. (1999). *Test Theory: a Unified Treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Neath, A. A., & Cavanaugh, J. E. (2011). The Bayesian information criterion: background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(2), 199–203. doi:10.1002/wics.199
- Raykov T. (1997). Scale reliability, cronbach's coefficient alpha, and violations of essential tau- equivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32, 329–353. doi:10.1207/s15327906mbr3204_2
- Schwarz G. (1978). Estimating the dimension of a model. *Annual Statistics*, 6, 461–464.
- Sijtsma K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika* 74, 107–120. 10.1007/s11336- 008-9101-0
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16(2), 209–220. doi:https://dx.doi. org/10.1037/a0023353