EC330
Tratamento da Informação

Capítulo sobre análise fatorial
Jacob A. Laros

Prof. Dr. Maurício U. Kleinke

PECIM – PPG em Ensino de Ciências e Matemática

Grupo de Ensino e Avaliação

DFA/IFGW/Unicamp

O Uso da Análise Fatorial: Algumas Diretrizes para Pesquisadores

Jacob A. Laros, PhD, UnB

IFGW
Instituto de Física Gleb Wataghin

Principais usos da análise fatorial

- Reagrupar uma gande quantidade de itens utilizados para medir personalidade, estilos de comportamento ou atitudes.

- Verificação da unidimensionalidade, analisar se o IA está de fato medindo mais do que um fator ou não.

Criação do fatores

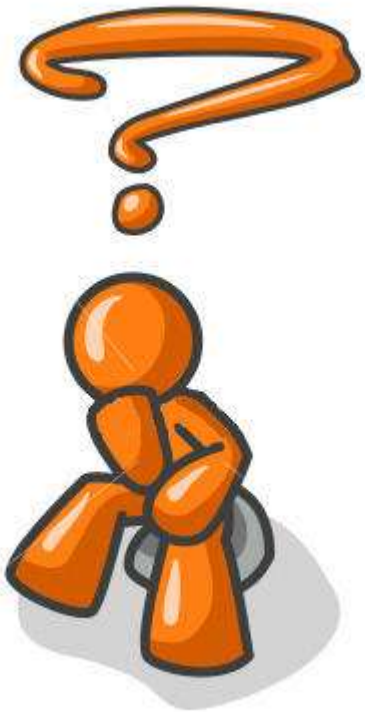
- Transformar um grande número de variáveis observadas em um número menor de variáveis hipotéticas, não-observadas, também chamadas de fatores, são responsáveis pelo fato de as variáveis observadas correlacionarem entre si.
- Poderoso método para reduzir a complexidade de uma grande quantidade de variáveis a uma estrutura relativamente simples, consistindo de um número menor de fatores. Os fatores são combinações lineares de variáveis observadas.

Validade fatorial e de constructo

- Os questionários se constituem em uma visão do pesquisador, do analista. A análise fatorial permite que essa visão seja confrontada com a visão dos avaliados, dos participantes da pesquisa, a partir dos resultados de construtos (fatores) surgidos durante a análise.
- Análise fatorial e validade de construto têm sido associadas durante muito tempo. A este respeito, Nunnally fez a seguinte observação: "*Análise fatorial é intimamente relacionada com as questões de validade de instrumentos psicológicos. Análise fatorial está no coração da mensuração dos construtos psicológicos*".

Técnica pouco compreendida

- Apesar da popularidade da análise fatorial, tanto a complexidade matemática dos procedimentos analíticos quanto a flexibilidade da técnica contribuíram para que esta, hoje em dia, seja uma das técnicas menos entendidas por psicólogos e estudantes de psicologia.



Pontos importantes da análise fatorial

1. A natureza e o tamanho da amostra que formará a base de dados da análise fatorial;
2. A adequação das distribuições das variáveis que entram numa análise fatorial;
3. A seleção de variáveis a serem submetidas à análise fatorial;
4. O uso de correlações corrigidas por atenuação ou o uso de correlações não corrigidas;
5. O tipo de correlações a ser utilizado;
6. A inclusão de variáveis marcadoras (*marker variables) entre as variáveis a serem analisadas;*

Pontos importantes da análise fatorial

7. *O número de fatores a serem extraídos;*
8. *A utilização da análise fatorial exploratória ou da análise fatorial confirmatória;*
9. *O tipo de análise fatorial a ser utilizado para extrair os fatores;*
10. *O procedimento de rotação a ser usado a fim de direcionar os fatores;*
11. *A interpretação dos resultados da análise fatorial;*
12. *A investigação de uma solução hierárquica;*
13. *A necessidade e a forma de calcular os escores fatoriais;*

Pontos importantes da análise fatorial

14. *A necessidade e a forma de calcular os escores fatoriais;*
15. *A investigação de validade cruzada da solução fatorial encontrada;*
16. *A utilização de grupos no estudo de invariância da estrutura fatorial encontrada;*
17. *A seleção dos resultados mais relevantes da análise fatorial para a publicação da pesquisa.*

Usos básicos e métodos de análise fatorial

Duas formas principais de utilizar a análise fatorial:

- Identificar as dimensões subjacentes de um determinado domínio (por exemplo: personalidade, criatividade, inteligência) que o instrumento em questão está medindo.
- A redução de dados busca descobrir ponderações ótimas para as variáveis mensuradas, de forma que um grande conjunto de variáveis possa ser reduzido a um conjunto menor de índices sumários (fatores) que tenham máxima variabilidade e fidedignidade.

Questionário sobre criatividade

Criativo

Com imagens

Com expressão corporal

Na expressão escrita

Na expressão oral

Utiliza a Análise Fatorial Comum

Os fatores são constructos já conhecidos

Questionário sobre aprendizagem

Aprendizagem

Eu estudo na sala de aula

Eu tenho um lugar sossegado para estudar em casa

Minha mãe me acompanha ao fazer o dever de casa

O/A professor(a) esclarece minhas dúvidas.

.....

Utiliza a Análise de Componentes Principais

Os fatores são resultado da análise estatística

1 – Decisão sobre a natureza e o tamanho de amostra

Heterogeneidade da amostra:

- Quanto mais heterogênea uma amostra, em relação às variáveis que estão sendo mensuradas, mais altas as correlações encontradas entre os escores de teste (a variância tem uma relação positiva com a correlação).
- Portanto, ao se utilizar a técnica de análise fatorial, melhores resultados são obtidos com amostras heterogêneas do que com amostras homogêneas.

1 – Decisão sobre a natureza e o tamanho de amostra

Tamanho da amostra:

- 10 sujeitos por variável, com um mínimo de 100 sujeitos na amostra total.
- Pelo menos 5 participantes por variável e uma amostra total de pelo menos 200 sujeitos.
- A partir de um estudo tipo *Monte Carlo, os autores sugerem que o tamanho desejado de uma amostra depende do tamanho das cargas fatoriais obtidas. Com cargas fatoriais em torno de 0,80, obtêm-se soluções fatoriais altamente estáveis em amostras de 50 pessoas. Quando as cargas fatoriais estão ao redor de 0,40, porém, amostras de 300 a 400 sujeitos são necessárias para atingir soluções estáveis.*
- *Um mínimo de 100 sujeitos por fator medido.*

1 – Decisão sobre a natureza e o tamanho de amostra

Tamanho da amostra:

- *Amostras de 50 sujeitos como muito inferiores;*
- *De 100 sujeitos como inferiores;*
- *De 200 como razoáveis;*
- *De 300 como boas;*
- *De 500 como muito boas;*
- *De 1.000 ou mais como excelentes.*

2 – Decisão sobre a adequação das distribuições das variáveis observadas

- Identificar e excluir itens que possuem distribuições extremamente assimétricas e desequilibradas.
- Em escalas com um formato *verdadeiro-falso, estes são os itens que quase todo mundo (por exemplo, 95% ou mais) endossa ou nega; em escalas com um formato tipo Likert, estes são os itens para os quais a grande maioria dos respondentes escolhe a mesma opção (por exemplo, "concordo completamente").*

2 – Decisão sobre a adequação das distribuições das variáveis observadas

- Análises fatoriais apresentam resultados mais confiáveis e com menor risco de interpretação quando os dados do IA apresentam uma distribuição normal.

- Alguns autores podem discordar sobre a intensidade, mas todos concordam que é preferível uma distribuição normal.

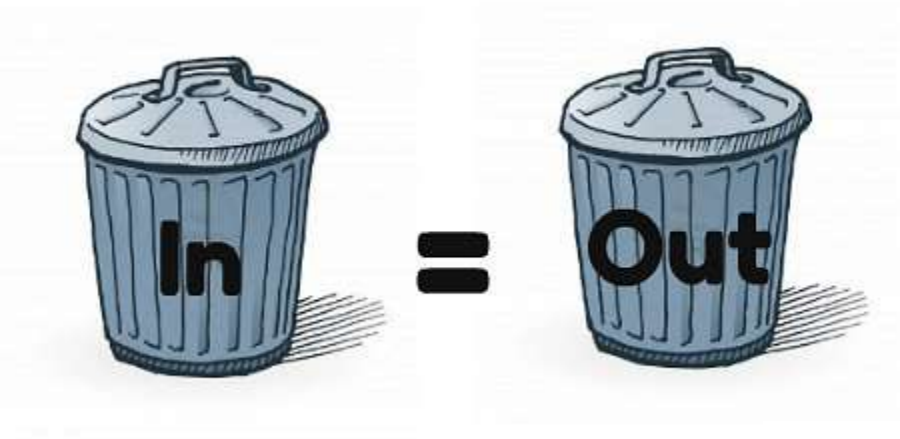
3 – Decisão a respeito da seleção das variáveis a serem submetidas à análise fatorial

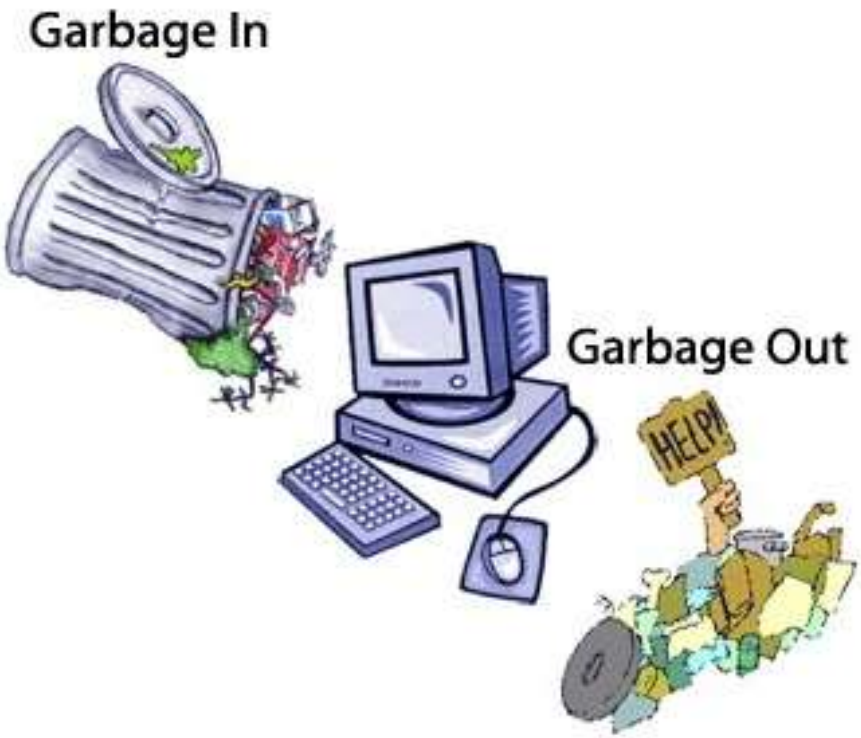
- Analisar as correlações entre os itens;
- Correlações abaixo de 0,30 são suspeitas!
- Itens dicotômicos podem apresentar problema na análise fatorial .

3 – Decisão a respeito da seleção das variáveis a serem submetidas à análise fatorial

- Os itens necessitam estar relacionados com seu constructo.
- Questões “bonitas” para banca elaboradora não servem para a prova!
- Questões muito “geniais” para questionários não serão compreendidas pelo seu público!

3 – Decisão a respeito da seleção das variáveis a serem submetidas à análise fatorial





4 – Decisão sobre a inclusão das variáveis marcadoras na análise fatorial

- Não é uma discussão muito próxima dos questionários que estamos pensando em aplicar ou utilizar;
- Depende de como um novo questionário se relaciona com os demais questionários que estão no mercado, principalmente na área de psicologia e de

5 – Decisão sobre o número de fatores a extrair

- Zwick e Velicer afirmaram que "*A determinação do número de componentes ou fatores a extrair é provavelmente a decisão mais importante que um investigador, executando análise fatorial, tomará*".
- *A ocorrência de erros nesta fase afetará a interpretação de todos os resultados subseqüentes.*
- *A superextração (a extração de um número maior de fatores dos que realmente existem) é um erro tão sério como a subextração (a extração de um número menor de fatores dos que realmente existem); ambos levarão a resultados e conclusões distorcidas.*

5 – Decisão sobre o número de fatores a extrair

(1) o critério de autovalor maior do que 1,0 de Guttman-Kaiser;

(2) o critério baseado no teste de qui-quadrado de Bartlett;

(3) o teste *scree de Cattell;*

(4) o critério da média mínima de correlações parciais de Velicer

(5) o critério de análise paralela de Horn.

5 – Decisão sobre o número de fatores a extrair

(1) o critério de autovalor maior do que 1,0 de Guttman-Kaiser;

(2) o critério baseado no teste de qui-quadrado de Bartlett;

(3) o teste *scree de Cattell;*

(4) o critério da média mínima de correlações parciais de Velicer

(5) o critério de análise paralela de Horn.

O critério de Guttman-Kaiser

- O critério de autovalor (*eigenvalue) maior do que 1,0 é o mais comumente utilizado, dado que é a opção padrão nos pacotes estatísticos como SPSS, SAS e BMDP.*
- *O autovalor corresponde à quantidade da variância explicada por um componente, sendo que um autovalor igual a 1,0 representa a totalidade de porcentagem da variância explicada por uma única variável.*

O critério de Guttman-Kaiser

- *Gorsuch relatou que o número de componentes retidos pelo critério GK normalmente se encontra entra 1/3 e 1/6 do número de variáveis incluídas na matriz de correlação.*

- *Vocês lembram do número de perguntas/ itens que deveria ter cada um dos constructos presentes no IA? Percebem a coincidência?*

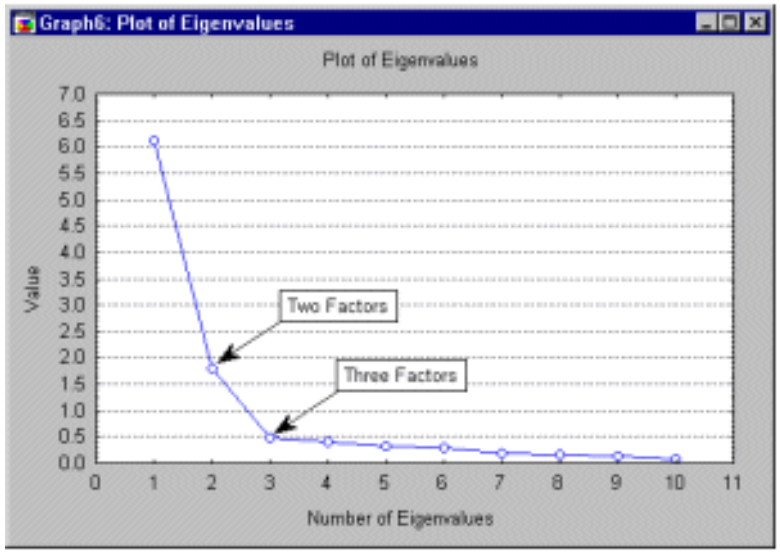
O teste *scree de Cattell*

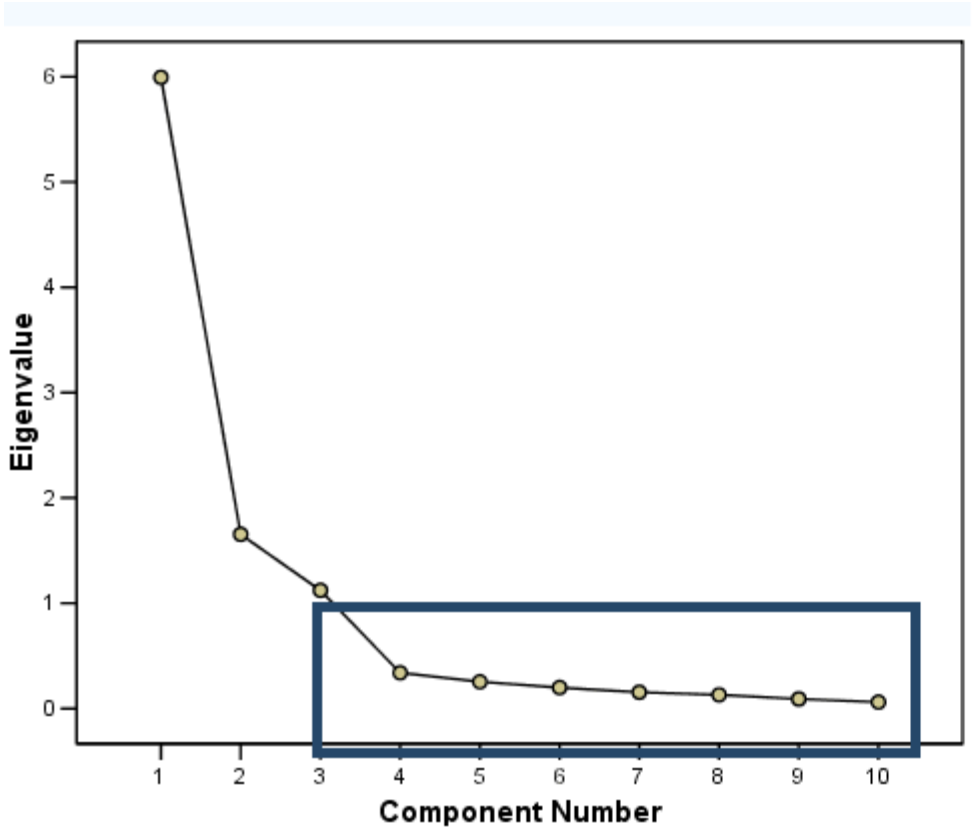
- *O critério é baseado em um gráfico dos autovalores. O eixo Y representa os valores dos autovalores e o eixo X mostra o número seqüencial dos componentes.*
- *O teste de scree separa o 'scree' de fatores triviais do início de fatores não triviais por intermédio de uma inspeção visual do gráfico. O julgamento subjetivo está então baseado no uso de uma linha reta colocada ao longo da parte do fundo do gráfico onde os pontos formam uma linha aproximadamente reta.*

O teste *scree de Cattell*

- *Os pontos acima da linha reta são associados com fatores não triviais, enquanto os pontos restantes representam os fatores triviais. Uma outra maneira para descrever o teste de scree consiste na busca de um "cotovelo'' no gráfico". Este cotovelo separa os fatores não triviais dos fatores triviais.*

O teste *scree de Cattell*





5 – Decisão sobre o número de fatores a extrair

Uma situação quali quanti

- Nem todos estão em todos os programas, o que dificulta a análise e a comparabilidade.
- Apesar de existirem vários indicadores para a escolha do número de fatores, considero que nada melhor que a análise qualitativa dos dados para estabelecer qual o melhor critério.

5 – Decisão sobre o número de fatores a extrair

Uma situação quali quanti

- Por análise qualitativa significa testar alguns números de fatores, e ver se os fatores que se obtém fazem sentido qualitativo, dentro da metodologia e dos pressupostos teóricos utilizados para a elaboração do IA.
- Em alguns pontos, o *feeling* do pesquisador é fundamental!

7 – Decisão sobre o procedimento de rotação

- Depois da extração, os fatores retidos geralmente são rotados para tornar a solução fatorial mais interpretável, mantendo as propriedades matemáticas da solução iguais.
- O objetivo do processo de rotação é conseguir uma estrutura fatorial simples.
- Uma estrutura simples é alcançada quando cada variável, preferencialmente, tem uma única carga alta em um único fator.

7 – Decisão sobre o procedimento de rotação

- Em outras palavras, uma estrutura fatorial simples existe quando cada variável tem uma carga principal em um único fator.
- A interpretação dos fatores, para a maioria dos pesquisadores, é simplificada quando todas as cargas fatoriais são positivas.

7 – Decisão sobre o procedimento de rotação

- Como obter somente alternativas positivas?

- Por exemplo, numa escala Likert de 5 alternativas, os valores 1 e 2 podem ser convertidos para os valores 4 e 5 e vice versa; o valor 3 que representa o ponto neutro da escala, permanece como está.

Cargas fatoriais e estrutura simples

Rotated Factor Pattern				
		Factor1	Factor2	Factor3
IMPED	IMPED	0.93487	0.1516	0.01142
REGRA	REGRA	0.93272	0.25943	0.005
GOSTA	GOSTA	0.74031	0.32808	0.38086
JOGAR	JOGAR	0.65536	0.29687	0.01628
VERTV	VERTV	0.22287	0.82423	0.27202
ESTAD	ESTAD	0.22495	0.81595	0.06885
MESAS	MESAS	0.30804	0.74294	0.19044
TORCE	TORCE	0.44299	0.66675	0.21068
VERCP	VERCP	0.13425	0.34533	0.82948
APOCP	APOCP	0.13397	0.26671	0.81287
ATVCP	ATVCP	0.1454	0.44899	0.72297
NVTCP	NVTCP	0.08889	0.13897	-0.74739

13 – Decisão de investigar a invariância da estrutura fatorial em amostras diferentes

14 – Decisão sobre a seleção dos resultados na publicação

Quem define os constructos?

- Ao longo de anos foi reconhecido cada vez mais que nenhum método analítico, inclusive análise fatorial, nos dita como nossos construtos ou nossas teorias deveriam ser (veja Mulaik, 1994).
- Como Mulaik enfatizou: “somos nós que criamos significados para coisas decidindo como eles deveriam ser usados. Assim, nós deveríamos ver a loucura de supor que a análise fatorial nos ensinará o que é que é inteligência, ou o que é personalidade”.

Quem define os constructos?

- Tais estudiosos discutem que nossos dados podem informar nosso julgamento, mas que nós somos responsáveis pelo julgamento que temos que exercitar inevitavelmente, e nossas definições de construto devem ser fundamentadas na teoria e não nos dados.