

1998 Volume 3 Nº 2

# Psicologia

Teoria, investigação e prática

ESIA

Outubro 1998

ISSN  
0873-4976

# CONSIDERAÇÕES SOBRE A ANÁLISE FACTORIAL DE ITENS COM RESPOSTA DICOTÓMICA

**Ricardo Primi**

Faculdade de Ciências Humanas, Universidade São Francisco, Brasil (\*)

**Leandro S. Almeida**

Instituto de Educação e Psicologia, Universidade do Minho

**Resumo** – *Através da análise factorial a psicometria procura explicar a relação entre resultados de um conjunto de variáveis observadas em função de um número menor de variáveis latentes. Estas representariam as estruturas internas e, portanto, as "causas" das diferenças interindividuais nos resultados das variáveis observadas. Assim, é uma prática muito comum utilizar a análise factorial para verificar a validade de conceito de questionários de personalidade ou dos testes de inteligência. Uma escala é válida conceptualmente quando se pode verificar empiricamente que os itens estão a medir a dimensão psicológica (ou seja a variável latente) que, de acordo com a teoria, se supõe que eles meçam. Contudo, quando os itens são dicotómicos um cuidado especial deve ser tomado quanto ao método de cálculo dos coeficientes de correlação. Ao longo deste artigo ilustramos o problema em causa e apresentamos algumas formas para superar as dificuldades enunciadas.*

**PALAVRAS-CHAVE:** Correlação phi e tetracórica; Análise factorial; Factores de dificuldade; Variáveis dicotómicas

**KEY WORDS:** Phi and tetrachoric correlations; Factor analysis; Difficulty factors; Dichotomic variables

## INTRODUÇÃO

Como afirma Cattell (1973, p.6), nas fases iniciais do desenvolvimento das ciências em geral, observam-se esforços para definir a taxonomia ou estrutura dos seus fenómenos particulares. Assim, a Química procurou, inicialmente, definir os elementos constituintes da matéria antes de elaborar teorias gerais sobre fenómenos complexos. Essa orientação esteve presente nos estudos sobre a personalidade e sobre inteligência desenvolvidos pelos psicometristas na

*Toda a correspondência relativa a este artigo deverá ser enviada para: Ricardo Primi, Faculdade de Ciências Humanas, Universidade São Francisco, Rua Alexandre Rodrigues Barbosa, 45, Itatiba, São Paulo, Brasil, CEP 13250-000 e-mail: rprimi@mpc.com.br.*

PSICOLOGIA: TEORIA, INVESTIGAÇÃO E PRÁTICA, 1998, 3, 225-234 © Centro de Estudos em Educação e Psicologia, Universidade do Minho

(\*) – Trabalho financiado pelo CNPq

primeira metade deste século. Tais pesquisas procuravam definir quais seriam os traços naturais subjacentes ou quais eram as estruturas constituintes da personalidade e inteligência humana que, enquanto causas, poderiam explicar o comportamento observável.

O objectivo central destes estudos nas diferentes dimensões psicológicas era identificar quais seriam os traços básicos ou os factores responsáveis pelas diferenças individuais, definindo o seu número e estrutura organizativa. Para isto, a psicometria recorreu à *análise factorial*, um método estatístico que busca analisar estruturas em matrizes de covariância ou correlação buscando redefini-las através de um número menor de variáveis latentes. Segundo Johnson e Wichern “o propósito essencial da análise factorial é descrever, se possível, as covariâncias entre variáveis em termos de um número menor de variáveis aleatórias subjacentes inobserváveis chamadas factores” (1992, p. 396).

Os grandes teóricos da definição e medida da personalidade e da inteligência no seio da abordagem psicométrica da Psicologia recorriam a um conjunto alargado de medidas psicológicas em amostras também elas bastante heterogéneas. Nesta altura, se os resultados se apresentavam altamente correlacionados entre si, então podia-se inferir a existência de uma única variável latente responsável por estas inter-relações. Interessante assinalar que, assumindo-se as variáveis latentes para explicar as intercorrelações entre variáveis manifestas, certo que se tornava necessário estas últimas para se dar um significado psicológico às primeiras. Várias dificuldades e críticas à abordagem psicométrica em Psicologia assentaram nesta tautologia compreensiva.

Como se afirmou, através da análise factorial a psicometria procurou explicar a relação entre resultados de diferentes variáveis observadas em função de um número menor de variáveis latentes. Nota-se, portanto, que este método tenta modelar as *diferenças entre indivíduos*, ou seja, os desvios em relação à média entre os  $n$  sujeitos, nas  $p$  medidas efetuadas, colocando-os (os desvios) em função de um conjunto menor de componentes latentes. Estas, representariam as estruturas internas, ou traços subjacentes e, daí, as “causas” das diferenças interindividuais nos resultados das variáveis observadas.

Face ao exposto, a análise factorial tornou possível o estudo empírico das dimensões psicológicas internas. O investigador partia de um conjunto de variáveis observáveis, e através das inter-relações entre elas, tentava identificar as possíveis dimensões subjacentes a esses comportamentos. Saliente-se que este método é puramente correlacional, isto é, não implica em nenhum momento a manipulação experimental. Numa analogia interessante, Cattell explica o método da análise factorial: “O problema que durante muitos anos desconcertou os psicólogos era encontrar um método que deslindasse essas influências funcionalmente unitárias na floresta caótica do comportamento humano. Mas como é que, numa floresta tropical de facto decide o caçador se as manchas escuras que vê são dois ou três troncos apodrecidos ou um só jacaré? Ele fica à espera de movimento. Se eles se movem juntos – aparecem e desaparecem juntos – ele conclui por uma única estrutura. Da mesma forma, como John Stuart Mill observou na sua filosofia da ciência, o cientista deveria ter em mira a “variação concomitante” na busca de conceitos unitários” (1965, p. 55).

Na Psicologia, o recurso à análise factorial generalizou-se bastante. Trata-se de uma prática muito comum quando se pretende verificar a validade de conceito dos instrumentos usados (Anastasi, 1981). Uma escala é válida conceptualmente quando se pode verificar que os itens avaliam a dimensão psicológica (ou seja a variável latente) que, teoricamente, é suposto estarem a avaliar. Se se supõe, por exemplo, que a escala SSQ (Rielo, 1992), embora sendo formada por 100 itens, avalia quatro dimensões da personalidade, espera-se encontrar quatro conjuntos de itens com alta covariação já que eles estariam associados a estas quatro dimensões. Portanto, espera-se que a covariação entre os 100 itens possa ser explicada por uma estrutura bastante mais simples com apenas quatro factores. A análise factorial permitiria verificar esta estrutura.

### MEDIDAS DE RELAÇÃO EM ITENS DICOTÓMICOS

Quando os itens são dicotómicos em termos do formato das respectivas respostas, um cuidado especial deve ser tomado com relação ao método de cálculo dos coeficientes de correlação. O problema nestes casos ocorre porque a amplitude de variação dos coeficientes de correlação é afectada caso as variáveis a serem correlacionadas não apresentem uma distribuição homogénea e simétrica como demonstrou Ferguson (1941, 1959). Como exemplo, consideremos o quadro I, contendo um caso hipotético das respostas de dez sujeitos a quatro itens. Os quatro itens são heterogéneos quanto ao seu índice de dificuldade e o padrão de resposta é altamente consistente.

Sujeitos	Item 1	Item 2	Item 3	Item 4	Resultado
Suj 1	1	1	1	1	4
Suj 2	1	1	1	1	4
Suj 3	1	1	1	1	4
Suj 4	1	1	0	1	3
Suj 5	1	1	0	1	3
Suj 6	0	0	0	1	1
Suj 7	0	0	0	1	1
Suj 8	0	0	0	0	0
Suj 9	0	0	0	0	0
Suj 10	0	0	0	0	0
ID	0.5	0.5	0.3	0.7	

Quadro I – Exemplo de um padrão perfeitamente consistente de resposta

Como estes itens são dicotómicos eles poderiam ser correlacionados utilizando o coeficiente *Phi*:

$$r_{ij} = \frac{p_{ij} - p_i p_j}{\sqrt{p_i p_j q_i q_j}}$$

Onde,

$r_{ij}$  = correlação *Phi* entre o item *i* e *j*

$p_{ij}$  = probabilidade conjunta de acerto no item *i* e *j*

$p_i$  = probabilidade de acerto no item *i*

$p_j$  = probabilidade de acerto no item *j*

$q_i = 1 - p_i$  e  $q_j = 1 - p_j$

Então a correlação entre os itens 1 e 2,

$$r_{12} = \frac{0.5 - 0.5 \cdot 0.5}{\sqrt{0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5}}$$

E a correlação para os itens 3 e 4,

$$r_{12} = \frac{0.3 - 0.3 \cdot 0.7}{\sqrt{0.3 \cdot 0.7 \cdot 0.7 \cdot 0.3}}$$

Pode-se observar que a correlação entre os itens 1 e 2 assume o valor máximo, no entanto, a correlação entre os itens 3 e 4 assume o valor 0,43. Neste caso, a correlação não assume o valor máximo, pois os itens têm distribuições diferentes. Como afirma Ferguson "se o coeficiente de correlação é visto como uma medida de eficácia da predição, então a predição perfeita em ambas as direcções positiva e negativa só é possível quando as duas distribuições têm a mesma forma e são simétricas. Se uma variável é normalmente distribuída e a outra é rectangularmente distribuída, a predição perfeita de uma em função da outra não é possível e o coeficiente de correlação reflete este facto" (1959, p. 422).

Considerando os itens 3 e 4, que apresentam respectivamente ID's 0.3 e 0.7, observa-se, no quadro I, que os três sujeitos que acertaram o item 3 também acertaram o item 4. No entanto, 4 pessoas que acertaram o item 4 não acertaram o item 3. Em termos de previsibilidade, isto implica uma falha, já que não é possível prever totalmente o desempenho no item 4 a partir do item 3, e vice versa. Se uma pessoa acertar o item 3, que é mais difícil, certamente acertará o item 4, que é mais fácil. No entanto se uma pessoa acertar o item 4, mais fácil, ela tanto poderá acertar o item 3, mais difícil, como ocorreu no caso dos sujeitos 1, 2 e 3, como errar (sujeitos 4, 5, 6 e 7). Coerente com este facto, o coeficiente de correlação não assume o seu valor máximo quando calculado para os itens 3 e 4.

Geralmente, falando-se de testes psicológicos, os autores têm procurado deliberadamente incluir itens com uma ampla variação de dificuldade. O objectivo deste procedimento é aumentar a sensibilidade dos resultados na prova através de uma adequada diferenciação dos desempenhos individuais (Almeida & Freire, 1997). Portanto, mesmo que haja um padrão de respostas altamente consistente, como o do quadro I, as intercorrelações podem ser baixas por causa das distribuições heterogéneas. Sendo assim, é necessário buscar uma forma em que o critério de previsibilidade possa ser "relaxado" para se adaptar as necessidades específicas destes casos.

Suponha-se, hipoteticamente, que todos os itens exemplificados acima avaliam um único traço de personalidade a saber introversão-extroversão (resultados mais elevados reflectiriam extroversão). Suponha-se, também, que esse traço seja uma variável contínua distribuída normalmente, e que as pessoas assumiam um determinado valor entre estes dois extremos. Suponha, por último, que os itens 3 e 4 fossem os seguintes (medindo a dimensão extroversão):

*Item 3. Quando vejo novos colegas na minha sala de aula*

a. eu me apresento imediatamente a eles (1)

b. eu me retraio e os observo por um tempo (0)

*Item 4. Na escola eu gosto de trabalhar*

a. em grupo (1)

b. individualmente (0)

O valor de ID=0,3 para o item 3 (ou seja menos respostas de valor 1) pode estar associado ao facto de que neste item só pessoas com alto grau de extroversão escolheriam a resposta "a". No item 4, no entanto, um grau menor de extroversão seria já suficiente para o sujeito responder "a". Esta situação encontra-se exemplificada na figura 1 através das duas curvas características do item e de uma curva normal representando a distribuição dos resultados na variável contínua latente (o valor das abcissas nas curvas características dos itens representam a probabilidade de escolher a resposta "1"). Considerando-se uma mesma probabilidade de escolha ( $p=0,5$ ), nota-se que para o item 1, um baixo valor de *theta* (que representaria extroversão) está associado a esta probabilidade (-1,6). Já no item 2, um valor mais alto de extroversão seria necessário para atingir esta probabilidade (+1,6).

Cronbach faz uma analogia interessante a propósito do conceito de probabilidade de "passar" um item em função de uma variável: "Numa escala 'pura', tarefas ou perguntas de dificuldade variável medem a mesma dimensão. Uma maneira de descobrir quantos gramas pesa o objecto X é colocá-lo no prato esquerdo de uma balança e colocar objectos de peso conhecido - 200g, 250g e assim por diante - no prato da direita. Cada um deles é como um item do teste. O objecto X "passa" no item quando o prato da esquerda desce, o teste continua até chegarmos a um peso suficientemente grande para fazer descer o lado direito da balança" (1996, p. 132).

A curva característica do item demonstra que, quanto maior for a extroversão do sujeito, maior será a sua probabilidade de "passar" no item. De acordo com a analogia acima, o item 3, que seria representado por um objecto colocado no lado direito da balança, teria um peso

ou refere uma intensidade de extroversão maior do que o item 4. O sujeito representado pelo objecto X, colocado no prato esquerdo, precisaria ser mais "pesado" que o item 3 para passá-lo, demonstrando alta extroversão.

A curva normal sobreposta revela que a probabilidade de existirem pessoas com valores extremos (muito extrovertidas ou muito introvertidas) é menor do que valores médios na variável latente. Portanto, há uma baixa probabilidade de existirem pessoas com alto grau de extroversão. Em função disso quando trabalhamos com itens que medem um mesmo traço com diferentes intensidades as distribuições de resposta serão diferentes.

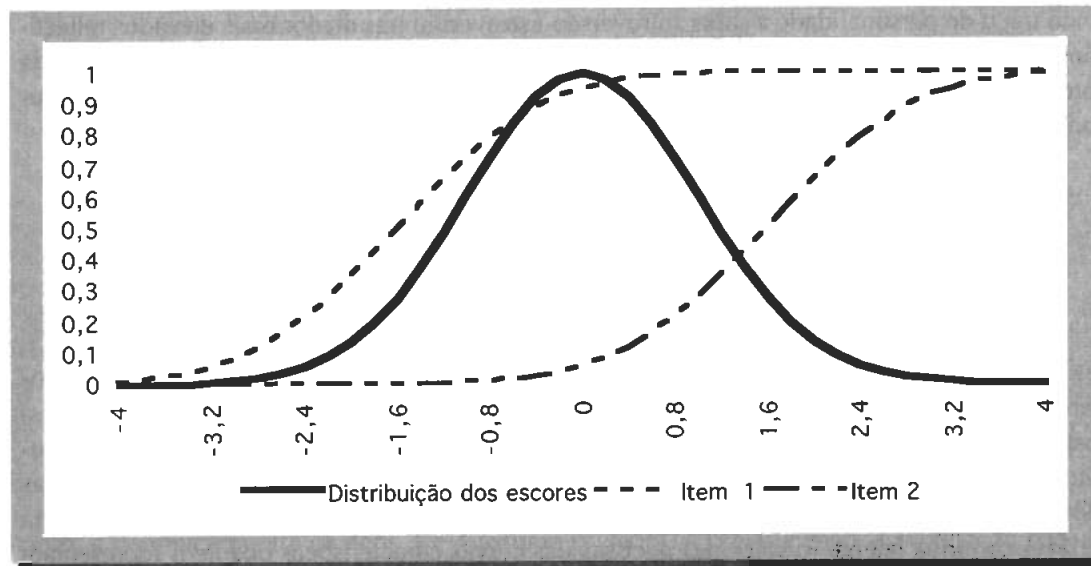


Figura 1 - Exemplo de dois itens, medindo um mesmo factor, com dificuldades diferentes

### ANÁLISE FACTORIAL DE ITENS COM DIFERENTES ID'S

Supondo, agora, que os itens medem uma mesma variável com dificuldades diferentes, espera-se que todos os que acertem os mais difíceis também acertem os mais fáceis. No exemplo acima, os sujeitos que estabelecem logo de início um primeiro contacto (item 3), deveriam também escolher actividades em grupo (item 4). Neste sentido, se todos os sujeitos que respondem 1 ao item 3 respondessem 1 ao item 4 estaríamos aptos a supor uma relação máxima entre os dois itens. Isto, no entanto, não ocorre se usarmos o coeficiente *Phi*.

O quadro II apresenta propositadamente estas propriedades, ou seja, os sujeitos com resultados mais fracos acertam os itens mais fáceis, enquanto os sujeitos com resultados superiores acertam os itens mais difíceis assim como os mais fáceis. O padrão de covariação entre os itens deve ser descrito por um único factor já que todos os itens mostram uma relação máxima,

dadas as restrições geradas por distribuições heterogêneas. Contudo a análise factorial destes itens, se baseada na matriz de correlações *phi*, explicaria as variações dos itens através de dois factores (quadro II).

	ID	Factor 1 Auto valor = 3,05 (76,4%)	Factor 2 Auto valor = 0,57 (14,3%)
item 1	0,5	0,96	0
item 2	0,5	0,96	0
item 3	0,3	0,77	-0,53
item 4	0,7	0,77	0,53

Quadro II - Resultados da análise factorial dos itens do quadro I

Ferguson (1941, p. 326) descreve este problema mostrando que o primeiro factor representa o factor geral esperado, contudo, o segundo factor representa uma dicotomização onde cada carga factorial representa o desvio da dificuldade que o item em causa tem com a média das dificuldades de todos os itens  $((0,5 + 0,5 + 0,3 + 0,7)/4 = 0,5)$ . Assim, este factor não teria significado em termos de unidades funcionais unitárias subjacentes já que a variação neste factor se associa a uma variação de dificuldade nos itens, não sendo possível atribuir significado psicológico a este factor.

Para se evitar este problema importa fazer com que os coeficientes de correlação assumam o valor máximo independentemente das distribuições. Sugerimos dois procedimentos para este efeito:

- (1) dividir o coeficiente *phi* pelo valor máximo que este pode assumir dada as restrições de distribuição. Segundo Ferguson (1941, p.324), o valor máximo possível é dado por:

$$r_{ij} = \sqrt{\frac{p_j q_i}{p_i q_j}}$$

Onde,

$r_{ij}$  = valor máximo que o coeficiente *Phi*, entre o item *i* e *j*, pode assumir dadas as distribuições heterogêneas

$p_i$  = probabilidade de acerto no item *i* (sendo *i* o item com maior probabilidade de acerto)

$p_j$  = probabilidade de acerto no item *j* (sendo *j* o item com menor probabilidade de acerto)

$$q_i = 1 - p_i \text{ e } q_j = 1 - p_j$$

(2) usar o coeficiente de correlação *tetracórico*. Supondo que na tabela 2X2 usada para cruzar os valores dicotómicos dos dois itens tenha as concordâncias notadas por A e D e as discordâncias por B e C como tradicionalmente se considera, um valor aproximado deste coeficiente é dado por (Carlson, 1995):

$$r_t = \cos \left( \frac{180^\circ}{1 + \sqrt{\frac{BC}{AD}}} \right)$$

O coeficiente *tetracórico* pressupõe que as duas variáveis a correlacionar são na sua natureza contínuas com distribuição normal e foram artificialmente dicotomizadas. Sendo assim, as diferenças de distribuição entre as variáveis dicotomizadas não interfere no seu valor dado que as duas terão a mesma distribuição. Dada esta suposição, o coeficiente de correlação *tetracórico* pode ser usado pois não terá seu valor máximo afectado pelas diferenças de distribuição.

Um cálculo mais exacto desse coeficiente envolve procedimentos interactivos mais complexos. O programa "Statistica for Windows" permite o cálculo deste coeficiente através do procedimento "Reliability Analysis". Ainda no procedimento "Prelis" do programa LISREL, obtêm-se matrizes de correlação com coeficientes *tetracóricos* (como aliás coeficientes *policóricos* utilizados quando se tem variáveis politómicas, por exemplo de formato "likert").

No quadro III, apresenta-se uma comparação dos valores obtidos através do coeficiente *phi* (sem correcção) e do coeficiente *tetracórico*. Esta comparação foi feita com base nos dados obtidos por Primi, Almeida e Lucarelli (1996) em relação às provas AR e NR da Bateria de Provas de Raciocínio Diferencial. Os dados indicam claramente que, se o coeficiente *phi* fosse utilizado, o valor-próprio do primeiro factor seria subestimado. Por este facto, sugere-se a utilização de um coeficiente *phi* corrigido (dividido pelo máximo possível) ou o coeficiente *tetracórico*.

Prova	Coefficiente	Média das inter-correlações entre os itens	Valor-próprio do primeiro factor
Prova AR (N = 213)	Phi	0.09	4.54 (12.98%)
	Tetracórico	0.73	11.71 (33.46%)
Prova NR (N = 204)	Phi	0.20	7.31 (24%)
	Tetracórico	0.38	12.15 (40.52%)

Quadro III - Comparação entre os coeficientes *phi* e *tetracórico* aplicados à análise factorial das provas AR e NR da BPRD.

## CONCLUSÃO

Vários instrumentos de avaliação psicológica recorrem a itens dicotómicos com diferentes índices de dificuldade na sua formulação ou com diferentes proporções de acordo. Esta situação é particularmente frequente nos testes de inteligência e aptidões, por exemplo, nos itens com várias alternativas de escolha (uma apenas correcta). Esta situação levanta algumas dificuldades quando os autores pretendem analisar a estrutura factorial subjacente aos desempenhos nos itens dessas provas.

Vários exemplos de tais dificuldades foram enunciadas ao longo deste artigo. Com base na literatura, foram sugeridos dois procedimentos alternativos para cálculo dos coeficientes de correlação em itens dicotómicos. A análise factorial de itens dicotómicos pode recorrer ao coeficiente de correlação *tetracórico* ou ao coeficiente *phi*. Neste último caso, divide-se o valor do coeficiente observado pelo valor máximo que este pode assumir dado pela fórmula apresentada. Como ficou demonstrado, o uso de coeficientes *phi* sem correcção subestimam o valor da associação entre as variáveis, afectando logicamente os parâmetros estatísticos dos factores isolados.

## REFERÊNCIAS

- Anastasi, A. (1961). *Testes psicológicos*. São Paulo: EPU.
- Carlson, R. (1995). *Correlation coefficients*. Texas: Texas Pan America University.
- Cattell, R. B. (1965). *Análise científica da personalidade*. São Paulo: Ibrasa.
- Cattell, R. B. (1956). Validation and intensification of the sixteen personality factor questionnaire. *Journal of Clinical Psychology*, 12, 205-214.
- Cattell, R. B. (1944). A note on correlation clusters and cluster search methods. *Psychometrika*, 9(3), 169-184.
- Cronbach, L. J. (1996). *Fundamentos da testagem psicológica*. Porto Alegre: Artes Médicas.
- Ferguson, G. A. (1959). *Statistical analysis in psychology and education*. New York: McGraw-Hill.
- Ferguson, G. A. (1941). The factorial interpretation of test difficulty. *Psychometrika*, 6 (5), 323-329.
- Johnson, R. A. & Wichern, D. W. (1992). *Applied multivariate statistical analysis*. London: Prentice Hall.
- Primi, R., Almeida, L. S. & Lucarelli, M. D. M. (1996). Bateria de Provas de Raciocínio Diferencial (BPRD): Resultados numa amostra de adolescentes de São Paulo. *Avaliação Psicológica: Formas e Contextos*, IV, 385-392.
- Riello, I. C. (1992). *O temperamento de adolescentes: Estudo de um instrumento de avaliação*. Dissertação

- de Mestrado. Campinas, São Paulo: Universidade Católica de Campinas.
- Takechi, S. (1996). *Métodos quantitativos de análise de dados observacionais*. São Paulo: Universidade de São Paulo, Departamento de Psicologia Experimental.
- Wherry, R. J. & Gaylord, R. H. (1944). Factor pattern of test items and tests as a function of the correlation coefficient: Content, difficulty, and constant error factors. *Psychometrika*, 9(4), 237-244.

## ISSUES REGARDING FACTOR ANALYSIS WITH DICHOTOMIC ITEMS

### ABSTRACT

Using factor analysis psychometry has tried to explain the relationship between a set of observable variables by a fewer number of latent factors. These would represent internal structure which can be thought as underlying causes of individual differences. It is a common practice to

*use factor analysis to verify the personality and intelligence tests construct validity. The construct validity is inferred when it is possible to verify the communality among items which had been previously stated theoretically. Although when items are dichotomously scored a special consideration must be considered in the way these associations are estimated. In this paper we have discussed this issue and showed some alternative approaches to deal with these problems.*