

# *Visão além do alcance: uma introdução à análise fatorial*

---

---

**Dalson Brito Figueiredo Filho**

**José Alexandre da Silva Júnior**

Programa de Doutorado em Ciência Política  
Universidade Federal de Pernambuco

---

---

**Resumo:** Como mensurar fenômenos que não podem ser diretamente observados? O principal objetivo desse artigo é demonstrar por que a análise fatorial é a resposta mais adequada para responder a essa pergunta. Metodologicamente, utilizamos um banco de dados com diferentes medidas de democracia para ilustrar como a técnica de análise fatorial de componentes principais pode ser utilizada para medir as duas dimensões da poliarquia propostas por Robert Dahl (1971): contestação e inclusividade. Em termos substantivos, esperamos facilitar a compreensão dessa técnica nas Ciências Sociais em geral e na Ciência Política em particular.

**Palavras-chave:** análise fatorial; democracia; poliarquia; mensuração; métodos quantitativos

**Abstract:** How do we measure phenomena that cannot be directly observed? The principal aim of this paper is to demonstrate why factor analysis technique is the best answer to this question. Methodologically, we use a database with different indicators of democracy to show how principal component analysis can be employed to measure the two polyarchy dimensions proposed by Robert Dahl (1971): contestation and inclusiveness. On substantive grounds, we hope to facilitate the understanding of factor analysis technique in Social Sciences in general and in Political Science in particular.

**Keywords:** factor analysis; democracy; polyarchy; measurement; quantitative methods

*“What is important to study cannot be measured and what can be measured is not important to study” (CONVERSE, 1964)*

*“Concepts without percepts are empty (...) percepts without concepts are blind” (KANT)*

## **Introdução**

Faz parte da rotina das Ciências Sociais mensurar fenômenos que não podem ser diretamente observados (BLALOCK, 1974; BOLLEN e GRANDJEAN, 1981; FIELD, 2005). Para Blalock, “ainda que o desenvolvimento da teoria seja importante em si mesmo, eu acredito que os mais sérios e importantes problemas que requerem nossa imediata e forte atenção são aqueles de conceitualização e mensuração” (BLALOCK, 1979: 882). Nesse sentido, um dos principais desafios enfrentados pelos pesquisadores é operacionalizar conceitos abstratos em variáveis empiricamente observáveis (BLALOCK, 1984). De acordo com Zeller e Carmines (1980), a transformação de conceitos em indicadores empíricos não é um processo simples. Pois o mesmo conceito pode ser operacionalizado de formas diferentes. Suponha que um pesquisador esteja interessado em medir vulnerabilidade social. Uma das possibilidades para tentar capturar esse conceito é eleger um *proxy*. Por exemplo, alguns estudos escolhem uma linha de pobreza específica como indicador de vulnerabilidade. Operacionalmente, o pesquisador pode escolher o *proxy* como representativo do seu conceito/variável de interesse, analisar como ele se relaciona com outras variáveis e utilizá-lo como variável dependente ou independente em estágios mais avançados de seu desenho de pesquisa.

Um processo alternativo de mensuração é identificar variáveis que “caminham juntas”, ou seja, variáveis que apresentam a mesma estrutura subjacente (TABACHINICK e FIDELL, 2007). Tecnicamente, isso pode ser implementado através da análise fatorial. A principal função das diferentes técnicas de análise fatorial é reduzir uma grande quantidade de variáveis observadas a um número reduzido de fatores. Os fatores representam as dimensões latentes (construtos) que resumem ou explicam o conjunto de variáveis observadas (HAIR et al, 2005). Ao resumir dados, a análise fatorial obtém dimensões latentes que descrevem os dados em um número menor de conceitos do que as variáveis individuais originais (HAIR et al, 2006, p. 91). Para Zeller e Carmines (1980), a análise fatorial não se refere a uma única técnica estatística, mas a uma variedade de técnicas relacionadas desenhadas para tornar os dados observados mais facilmente interpretáveis. Em resumo, existem ao menos duas formas de mensurar variáveis que não são empiricamente observáveis: escolher um *proxy* ou reduzir um

conjunto de variáveis a poucas dimensões através de análise fatorial. Embora mais simples, o primeiro caminho pode ser mais arriscado já que em muitas oportunidades não é possível identificar *proxies* adequados para representar os conceitos de interesse. Por outro lado, a análise fatorial oferece uma alternativa que agrega mais informações, no entanto, tende a exigir do pesquisador um maior grau de sofisticação técnica. O principal objetivo desse artigo é apresentar a lógica da análise fatorial a partir de uma abordagem intuitiva.

Não sem razão, uma motivação adicional que orienta este trabalho é a tímida utilização dessa técnica nas Ciências Sociais brasileira. Acreditamos que esse fenômeno pode ser explicado por dois principais componentes: (1) a resistência dos cientistas sociais brasileiros aos métodos quantitativos (SOARES, 2005; WERNECK VIANNA et al 1988; VALLE e SILVA, 1999; SANTOS e COUTINHO, 2000) e (2) o grau de complexidade matemática envolvido na operacionalização das diferentes técnicas de análise fatorial (DUNTEMAN, 1989). Um entrave adicional à utilização das técnicas de análise fatorial é que os benefícios associados à sua utilização ainda não são evidentes do ponto de vista conceitual (MOONEY, 1996). Dessa forma, enquanto não ficarem claras as suas potencialidades, é improvável que esse repertório de técnicas seja incorporado ao cotidiano dos pesquisadores brasileiros.

Portanto, nossa questão de pesquisa é: como diferentes variáveis podem ser resumidas em um conjunto reduzido de fatores? Metodologicamente, utilizamos um banco de dados com diferentes medidas de democracia para ilustrar como a análise fatorial de componentes principais pode ser utilizada para mensurar as duas dimensões da poliarquia propostas por Robert Dahl (1971): contestação e inclusividade. Teoricamente, essa análise foi pioneiramente realizada por Coppedge, Alvarez e Maldonado (2008). Similarmente, Baker e Koesel (2001) utilizam o conceito de poliarquia “*plus*” mensurada a partir da análise fatorial de 20 indicadores de democracia para países do leste europeu (1992-2000). Aqui nosso foco refere-se mais ao *modus operandi* da técnica do que à interpretação substantiva dos resultados, nosso interesse é apresentar cada passo da análise fatorial bem como as vantagens conceituais associadas à sua utilização.

Para tanto, o artigo está dividido em cinco partes. A primeira seção revisa brevemente parte da literatura sobre a análise fatorial; a segunda sumariza os principais requisitos que devem ser satisfeitos para utilizar essa técnica. A terceira parte oferece um exemplo de um desenho de pesquisa utilizando análise fatorial de componentes principais. Depois disso, descrevemos as principais estatísticas de interesse bem como a sua respectiva interpretação. A quinta seção apresenta as principais conclusões do artigo.

## Breve revisão da literatura<sup>1</sup>

De acordo com BARTHOLOMEW (1984), “a análise fatorial é uma técnica amplamente estatística, mas seus fundamentos teóricos são um tanto obscuros e sujeitos a disputa” (BARTHOLOMEW, 1984, p. 221). No entanto, atribui-se aos trabalhos de Charles Spearman e Karl Pearson os primeiros passos da análise fatorial (ZELLER e CARMINES, 1980). Spearman (1904) testou a hipótese de que diferentes testes de habilidade mental - habilidades em matemática, verbais, raciocínio lógico, entre outras - poderiam ser explicadas por um fator comum de inteligência que ele denominou “g”. De acordo com Kaplunovsky (2009), outra contribuição relevante foi feita por Thurstone (1935) ao desenvolver a idéia de *multiple factor analysis*. Outras contribuições significativas podem ser creditadas a Hotelling ao propor “o método de componentes principais que permite o cálculo da única matriz de fatores ortogonais” (KAPLUNOVSKY, 2009). Dessa forma, o desenvolvimento das diferentes técnicas de análise fatorial pode ser explicado a partir do acúmulo de diferentes trabalhos.

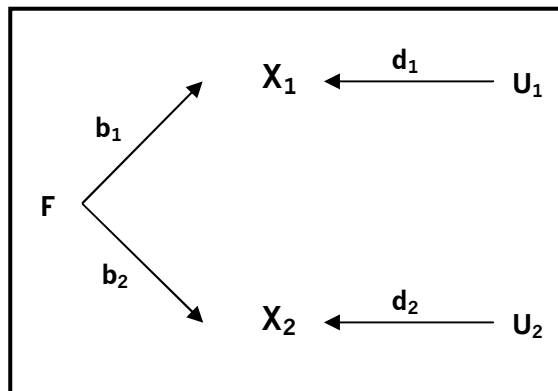
Mas para que serve a análise fatorial? De acordo com King (2001) “no modelo da análise fatorial, há muitas variáveis observadas cujo objetivo é gerar fatores subjacentes não observados (KING, 2001, p. 682). Ou seja, a principal função das diferentes técnicas de análise fatorial é reduzir uma grande quantidade de variáveis observadas em um número menor de fatores. Mas o que são fatores? Hair et al (2005) definem fator como a combinação linear das variáveis (estatísticas) originais<sup>2</sup>. A Figura 1 ilustra a relação entre variáveis diretamente observadas e os seus respectivos fatores.

---

<sup>1</sup> Para os propósitos desse artigo, o grau de complexidade matemática foi minimizado. Para os leitores interessados em aprofundar seus conhecimentos sugerimos cobrir a bibliografia citada. Para trabalhos clássicos sobre análise fatorial ver Harman (1967), Rummel (1967; 1970), Cooper (1983) e Bartholomew (1984). Para textos introdutórios ver Kim e Mueller (1978a; 1978b), Zeller e Carmines (1980), Decoster (1998) e Costelo e Orborne (2005). Para uma abordagem mais aprofundada ver Tabachnick e Fidell (2007), Lawley e Maxwell (1973), Isogawa e Okamoto (1980), Yalcin e Amemiya (2001) e Bollen e Arminger (1991). Para análise fatorial de dados *missing* ver Mackelprang (1970) e Ligny et al (1981), para análise fatorial de dados categóricos ver Bartholomew (1980) e Vermunt e Magidson (2004). Para aplicações práticas utilizando o SPSS ver Dancey e Reidy (2004), Pallant (2007) e Ho (2006). Para uma introdução em português ver Hair et al (2005). Para diferentes aplicações ver Grumm (1963), Roberts (1987), Dunn, Schneck e Lawson (1973), Haydyk et al (1995), Bonjean e Browning (1969) e Slatin (1974).

<sup>2</sup> Para Bartholomew (1984) “há às vezes debate sobre se essas variáveis latentes são reais em qualquer sentido, mas elas podem ser vistas simplesmente como construtos desenhados para simplificar e resumir a complexa rede de variáveis interrelacionadas com que a natureza nos confronta” (BARTHOLOMEW, 1984, p. 221).

**Figura 1**  
**Modelo das vias para duas variáveis, modelo de um fator comum**



Fonte: Asher, 1983

Na figura acima,  $X_1$  e  $X_2$  são variáveis observadas:  $X_1$  é causado por  $F$  e por  $U_1$ . da mesma forma,  $X_2$  é causado por  $F$  e por  $U_2$ . Na medida em que  $F$  é comum a  $X_1$  e  $X_2$  ele é considerado um fator comum. Contrariamente, tanto  $U_1$  quanto  $U_2$  são considerados fatores únicos já que são restritos a  $X_1$  e  $X_2$ , respectivamente (ASHER, 1983). De acordo com Kim e Mueller (1978), “a análise fatorial se baseia no pressuposto fundamental de que alguns fatores subjacentes, que são em menor número que as variáveis observadas, são responsáveis pela covariação entre as variáveis” (KIM e MUELLER, 1978, p. 12). Nesse exemplo  $F$ ,  $U_1$  e  $U_2$  são considerados fatores (não podem ser diretamente observados) enquanto que  $X_1$  e  $X_2$  são as variáveis que o pesquisador pode observar diretamente. É nesse sentido que “a análise fatorial supõe que as variáveis observadas (medidas) são combinações lineares de algumas variáveis-fonte subjacentes (fatores). Isto é, ela supõe a existência de um sistema de fatores subjacentes e um sistema de variáveis observadas” (KIM e MUELLER, 1978, p. 7-8)<sup>3</sup>.

Dito isso, é importante atentar para o fato de a literatura diferenciar duas principais modalidades de análise fatorial: exploratória e confirmatória (TABACHINICK e FIDELL, 2007). A análise fatorial exploratória (AFE) geralmente é utilizada nos estágios mais embrionários da pesquisa, no sentido de literalmente

<sup>3</sup> King (2001) adverte que “um erro comum consiste em ver as variáveis observadas como causas do fator. Isso é incorreto. O modelo correto tem variáveis dependentes observáveis como funções dos fatores subjacentes e não (KING, 2001, p. 682). Ou seja, por mais intuitivo que seja acreditar que as variáveis observadas causam o fator, a interpretação correta é justamente o posto: o fator é um construto (dimensão) comum entre as variáveis. Para a diferença na interpretação entre fatores e componentes ver Tabachinick e Fidell (2007).

explorar os dados. Nessa fase, procura-se explorar a relação entre um conjunto de variáveis, identificando padrões de correlação. Além disso, a AFE pode ser utilizada para criar variáveis independentes ou dependentes que podem ser utilizadas posteriormente em modelos de regressão. Por sua vez, a análise fatorial confirmatória (AFC) é utilizada para testar hipóteses. Nesse caso, o pesquisador guiado por alguma teoria testa em que medida determinadas variáveis são representativas de um conceito/dimensão<sup>4</sup>. Para os nossos propósitos, utilizaremos análise fatorial confirmatória (AFC) para demonstrar como diferentes indicadores de democracia podem ser sumarizados para representar as duas dimensões da poliarquia propostas por Dahl (1971): contestação e inclusividade.

### O planejamento de uma análise fatorial

Para iniciar o lado mais operacional da missão deste artigo é necessário observar que requisitos precisam ser satisfeitos para a utilização da análise fatorial (TABACHINICK e FIDELL, 2007). O objetivo dessa seção é sumarizar essas informações. A Tabela 1 sintetiza o planejamento de uma análise fatorial em três estágios.

**Tabela 1**  
**Planejamento da análise fatorial em três estágios**

Procedimento	O que deve ser observado
Verificar a adequabilidade da base de dados	Nível de mensuração das variáveis, tamanho da amostra, razão entre o número de casos e a quantidade de variáveis e o padrão de correlação entre as variáveis.
Determinar a técnica de extração e o número de fatores a serem extraídos	O tipo de extração ( <i>principal components, principal factors, image factoring; maximum likelihood factoring; alpha factoring; unweighted least squares; generalized least squares</i> ).
Decidir o tipo de rotação dos fatores	Se for ortogonal ( <i>Varimax, Quartimax, Equamax</i> ), se for oblíqua ( <i>direct oblimin, Promax</i> ).

<sup>4</sup> Uma utilização adicional da análise fatorial é em modelos de equações estruturais (*Structural Equation Modeling*). Para uma introdução ver Kline (2004) e Hair et al (2006). Para aplicações mais avançadas ver Bollen (1989, 1993, 2006). Para *softwares* ver Lisrel e AMOS. Para dois tutoriais na internet ver <<http://davidakenny.net/cm/causal.htm> e <http://www2.gsu.edu/~mkteer/semfaq.html>>

Caso este artigo fosse meramente teórico, nesta seção seríamos obrigados apenas a informar que: 1) todas as variáveis trabalhadas são contínuas; 2) o método de extração dos fatores é por componentes principais e 3) a rotação deles é feita pelo método ortogonal de *Varimax*. Mas, o que cada uma dessas escolhas implica, e que supostos estão por traz de cada uma delas? Como em outras técnicas, o ponto de partida é *verificar a adequabilidade da base de dados*. Em relação ao nível de mensuração, a literatura mais conservadora recomenda apenas a utilização de variáveis contínuas ou discretas. Hair et al (2006) aconselham evitar a inclusão de variáveis categóricas, e caso seja necessário, recomenda-se a inclusão de variáveis *dummies*. De toda forma, é importante que os procedimentos sejam devidamente reportados e as estatísticas sejam adequadamente registradas para que os leitores possam analisar criticamente o nível de confiabilidade dos resultados encontrados. Já King (2001) adverte que determinadas variáveis como sexo e cor nunca devem ser incluídas em um modelo de análise fatorial já que é improvável que algum fator influencie a sua variação. Dessa forma, além dos critérios técnicos é necessário considerar teoricamente como os fatores se relacionam com as variáveis observadas<sup>5</sup>.

Em relação ao número de casos, quanto maior, melhor. Hair et al (2006) sugerem que a amostra deve ser superior a 50 observações, sendo aconselhável no mínimo 100 casos para assegurar resultados mais robustos (HAIR et al, 2006). A razão entre o número de casos e a quantidade de variáveis deve exceder cinco para um ou mais (HAIR et al, 2006).

No que concerne ao padrão de correlação entre as variáveis, a matriz de correlações deve exibir a maior parte dos coeficientes com valor acima de 0,30. O teste de Kaiser-Meyer-Olkin (KMO) varia entre 0 e 1. Quanto mais perto de 1, tanto melhor. Palant (2007) sugere 0,6 como um limite razoável. Friel (2009) sugere a seguinte escala para interpretar o valor da estatística KMO: entre 0,90 e 1 excelente; entre 0,80 e 0,89 bom; entre 0,70 e 0,79 mediano; entre 0,60 e 0,69 medíocre; entre 0,50 e 0,59 ruim e entre 0 e 0,49 inadequado. Já Hair et al (2006) sugerem 0,50 como patamar aceitável. Por fim, a estatística *Bartlett Test of Sphericity* (BTS) deve ser estatisticamente significativa ( $p < 0,05$ ). A Tabela 2 a seguir sintetiza essas informações.

---

<sup>5</sup> Existem técnicas de estimação que permitem trabalhar com diferentes níveis de mensuração (KRITZER, 1978a; 1978b). Por exemplo, Vermunt e Magidson (2004) desenvolveram uma técnica chamada LCFA (*Latent Class Factor Analysis*) que permite trabalhar com variáveis categóricas e ordinais. Outra técnica é a BFA (*Binary Factor Analysis*) também conhecida como Análise Fatorial Boeliana (KREPT, 2004). O SPSS tem uma função denominada CATPCA - *Categorical Principal Component Analysis* - que permite trabalhar com variáveis categóricas. O pesquisador deve escolher as opções *Analyze*, *Data reduction*, *Optimal Scaling*.

**Tabela 2**  
**Síntese do primeiro estágio do planejamento da AF**

Nível de mensuração	Variáveis contínuas e discretas
Amostra	Amostras mínimas entre 50 e 100 casos; razão entre o número de observações e a quantidade de variáveis igual ou superior a cinco.
Correlação	Maior parte dos coeficientes de correlação deve apresentar valores acima de 0,30.
KMO	Quanto maior melhor, tendo 0,50 como o patamar mínimo de adequabilidade (Hair et al 2006)
BTS	$p < 0,05$

Depois de *verificar a adequabilidade da base de dados*, o pesquisador deve seguir para o segundo estágio: determinar a técnica de extração dos fatores (componentes principais, fatores principais, fatora  o por imagem; fatora  o por verossimilhan  a m  xima; fatora  o alfa; m  nimos quadrados n  o ponderados; m  nimos quadrados)<sup>6</sup>. Aqui vale destacar a diferen  a entre an  lise de componentes principais (ACP) versus an  lise fatorial (AF). Ambas as t  cnicas procuram produzir combina  es lineares de vari  veis que capturem o m  ximo poss  vel a vari  ncia das vari  veis observadas. Na ACP *toda* a vari  ncia    utilizada. Na AF apenas a vari  ncia *compartilhada* (DANCEY e REIDY, 2004). Tabachnick e Fidell argumentam que: “se voc   estiver interessado numa solu  o te  rica n  o contaminada por variabilidade de erro, a an  lise fatorial deve ser sua escolha. Se voc   quiser simplesmente um resumo emp  rico do conjunto de dados, a an  lise de componentes principais    uma escolha melhor” (TABACHINICK e FIDELL, 2007, p. 608). Para Garson (2009), “a an  lise de componentes principais    em geral preferida para fins de redu  o de dados (traduzindo o espa  o das vari  veis num espa  o   timo de fatores), enquanto a an  lise fatorial    em geral preferida quando o objetivo da pesquisa    detectar a estrutura dos dados ou a modelagem causal”. De acordo com Hair et al (2006), na maioria dos casos tanto a ACP, quanto a AF, chegam aos mesmos resultados se o n  mero de vari  veis superar 30 ou se as comunalidades excederem 0,60 para a maior parte das vari  veis. Este artigo utilizar   o m  todo de componentes principais por ser o mais utilizado (PALLANT, 2007).

Depois de decidir o m  todo de extra  o, o pesquisador deve determinar o n  mero de fatores que ser  o extra  dos. O objetivo da extra  o dos fatores (*factor*

<sup>6</sup> Todas essas formas de rota  o est  o dispon  veis na se  o *Extraction* do pacote estat  stico SPSS, vers  o 16.0. Para a diferen  a entre as diferentes t  cnicas de extra  o de fatores ver Tabachinick e Fidell (2007).



*extraction*) é determinar a quantidade de fatores que melhor representa o padrão de correlação entre as variáveis observadas. Aqui o pesquisador enfrenta um *trade off* entre parcimônia e explicação. Quanto mais fatores forem extraídos, menor é o grau de parcimônia, no entanto, maior é a quantidade total de variância pelos fatores. Por outro lado, quanto menos fatores forem extraídos, maior é o grau de parcimônia, todavia, menor será a quantidade total de variância carregada pelos fatores. Dessa forma, a solução ótima é identificar o número mínimo de fatores que maximiza a quantidade de variância total explicada.

Apesar de não existir um critério consensual para definir quantos fatores devem ser extraídos, a literatura aponta alguns métodos que podem auxiliar o pesquisador na hora de tomar essa decisão. Por exemplo, a regra do *eigenvalue*<sup>7</sup> (critério de Kaiser) sugere que devem ser extraídos apenas os fatores com valor de *eigenvalue* acima de um. Isso porque se o fator apresenta baixo *eigenvalue*, ele está contribuindo pouco para explicar a variância nas variáveis originais. Para Tabachnick e Fidell (2007), esse método funciona melhor quando o pesquisador utiliza entre 20 e 50 variáveis.

Um método auxiliar é o *Scree test* (Catell 1966; 1978). Para tanto, o pesquisador deve analisar graficamente a dispersão do número de fatores até que a curva da variância individual de cada fator se torne horizontal ou sofra uma queda abrupta. Em ambas as situações, isso indica que muita variância foi perdida e, por isso, deve-se parar de extrair fatores<sup>8</sup>. Adicionalmente, o pesquisador pode utilizar o critério da variância acumulada para determinar a quantidade de fatores que devem ser extraídos. Hair et al (2006) sugerem o patamar de 60% como sendo aceitável. Dessa forma, a extração dos fatores deve continuar até que o referido patamar seja alcançado. Por fim, no caso da análise fatorial confirmatória, além dos critérios estatísticos também é importante apresentar razões teóricas para justificar a extração dos fatores. Nesse sentido, o pesquisador deve justificar em termos conceituais qual é o padrão de relação esperado entre as variáveis observadas e os fatores.

---

<sup>7</sup> Para Garson (2009), “o *eigenvalue* de um dado fator mede a variância em todas as variáveis que é devida ao fator. A razão de *eigenvalues* é a razão da importância explicativa dos fatores em relação às variáveis. Se um fator tem um *eigenvalue* baixo ele contribui pouco para a explicação das variâncias nas variáveis e pode ser ignorado como redundante em relação a fatores mais importantes”.

<sup>8</sup> *Horn Parallel Analysis* (1965) é outro método de extração. Ele consiste em comparar os valores dos *eigenvalues* da amostra piloto com os valores gerados a partir de uma amostra aleatória do mesmo tamanho. Apenas são retidos os *eigenvalues* com valor superior aos gerados na amostra aleatória.

**Tabela 3**  
**Síntese do segundo estágio do planejamento da AF**

<b>Tipo de extração</b>	Determinar a técnica de extração dos fatores ( <i>principal components, principal factors, image factoring; maximum likelihood factoring; alpha factoring; unweight least square; generalized least square</i> )
<b>Regra de Kaiser</b>	Devem ser extraídos apenas os fatores com valor do <i>eigenvalue</i> acima de 1
<b>Scree test</b>	Analisar graficamente a dispersão do número de fatores até que curva da variância individual de cada fator se tornar horizontal ou sofrer uma queda abrupta
<b>Variância acumulada</b>	> 60%
<b>Razão teórica</b>	O pesquisador deve justificar teoricamente como as variáveis se relacionam com os fatores extraídos.

Depois de verificar a adequabilidade da base de dados e determinar a técnica de extração e o número dos fatores, o pesquisador deve seguir para o terceiro estágio: decidir o tipo de rotação dos fatores. De acordo com Schawb, “o método de rotação se refere ao método matemático que rotaciona os eixos no espaço geométrico. Isso torna mais fácil determinar quais variáveis são carregadas em quais componentes” (SCHAWB, 2007). Ou seja, o principal objetivo da rotação dos fatores é tornar o resultado empírico encontrado mais facilmente interpretável, conservando as suas propriedades estatísticas<sup>9</sup>. De acordo com Tabachinick e Fidell (2007), existem dois principais tipos de rotação: ortogonal e oblíqua. Por um lado, as rotações ortogonais são mais fáceis de reportar e de interpretar. No entanto, o pesquisador deve assumir que os construtos são independentes (na prática esse pressuposto é mais difícil de ser respeitado). Por outro lado, as rotações oblíquas permitem que os fatores sejam correlacionados. Todavia, são mais difíceis de descrever e interpretar. Em geral, as duas formas de rotação produzem resultados bastante semelhantes, principalmente quando o padrão de correlação entre as variáveis utilizadas é claro (PALLANT, 2007; HAIR et al, 2006). De acordo com Pallant (2007), o tipo de rotação ortogonal *Varimax* é o mais comumente utilizado (esse método procura minimizar o número de variáveis que apresentam altas cargas em cada fator). Por esse motivo, esse artigo utilizará esse tipo de rotação de fatores<sup>10</sup>.

<sup>9</sup> Para a diferença entre os diferentes tipos de rotação ver os Anexos do artigo e Tabachinick e Fidell (2007).

<sup>10</sup> Para a diferença detalhada das diferentes formas de rotação ver Tabachinick e Fidell (2007).

### **Exemplo de desenho de pesquisa: Dahl e a democracia**

Dahl argumenta que “uma característica chave da democracia é a contínua responsividade do governo às preferências dos cidadãos, considerados como politicamente iguais” (DAHL, 1971, p.25). Isso significa que os cidadãos devem ter oportunidade de: formular suas preferências; apresentar essas preferências a outros cidadãos e ao governo através da ação individual ou coletiva; ter essas preferências igualmente respeitadas na condução do governo, ou seja, sem discriminação pelo tipo ou fonte da preferência (DAHL, 1971, p. 02). Essas são as condições básicas para considerar um regime democrático. No entanto, para garantir que essas condições sejam efetivamente implementadas, o autor argumenta que é necessário assegurar oito garantias institucionais. A Tabela 4 sumariza essas informações.

**Tabela 4**  
**Alguns requisitos de uma democracia para um grande número de pessoas**

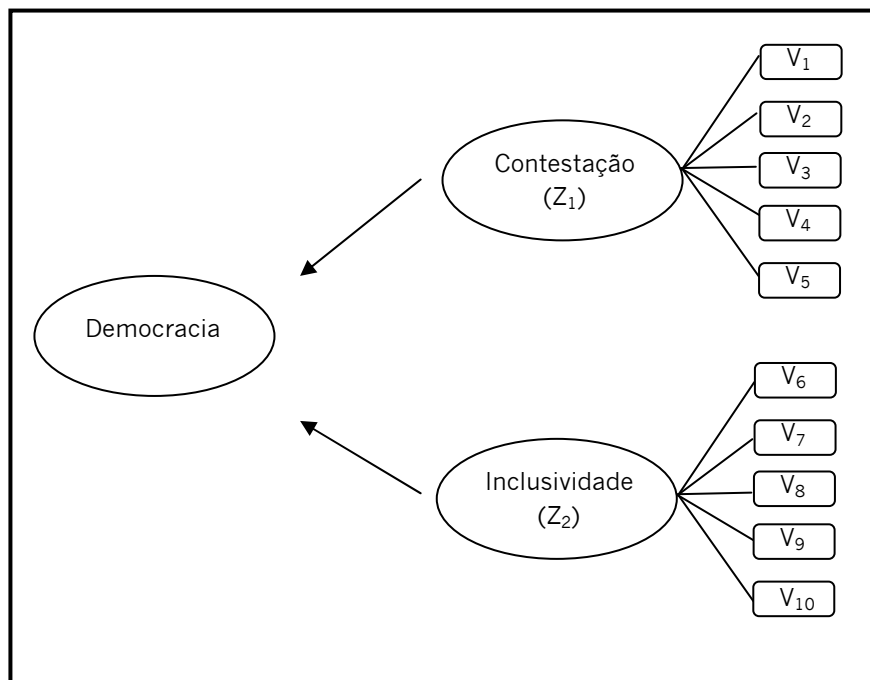
	<b>São necessárias as seguintes garantias institucionais:</b>
I. Formular preferências II. Exprimir preferências III. Ter preferências igualmente consideradas na conduta do governo.	1. Liberdade de formar e aderir a organizações; 2. Liberdade de expressão; 3. Direito de voto; 4. Elegibilidade para cargos públicos; 5. Direito de líderes políticos disputarem apoio; 6. Fontes alternativas de informação; 7. Eleições livres e idôneas; 8. Instituições para fazer com que as políticas governamentais dependam de eleições e de outras manifestações de preferência.

Fonte: reproduzido pelos autores a partir de Dahl (1971)

As oito garantias mencionadas permitem que os regimes sejam classificados em uma escala do menos ao mais democrático. Além disso, essas garantias podem ser reduzidas a duas dimensões: liberalização e inclusividade. Pelo primeiro, Dahl entende o reconhecimento do direito de contestação. A inclusividade diz respeito à ampliação do número de pessoas formalmente incorporadas ao processo político. Nesse sentido, quanto mais desenvolvidas forem essas duas dimensões, maior é o grau de democratização de um determinado

regime. É nesse sentido que o conceito de poliarquia proposto por Dahl (1971) constitui uma das mais influentes tentativas de traduzir um conceito em um indicador empírico. De acordo com Coppedge, Alvarez e Maldonado (2008) “cerca de três quartos do que a Polity, a Freedom House e outros indicadores têm medido consiste em variações das duas dimensões da democracia que Robert Dahl propôs em *Poliarquia – contestação e inclusividade*” (COPPEDGE, ALVAREZ e MALDONADO, 2008, p. 01). A Figura 2 ilustra essa lógica, onde  $V_1$  até  $V_{10}$  são variáveis observadas em diferentes bancos de dados sobre democracia<sup>11</sup>.

**Figura 2**  
**Dimensões de Dahl (1971) e variáveis observadas**



Nosso desenho de pesquisa replica o teste de hipótese efetuado por Coppedge, Alvarez e Maldonado (2008)<sup>12</sup>. A próxima seção apresenta as principais estatísticas de interesse e as suas respectivas interpretações.

<sup>11</sup> Para uma análise das vantagens e desvantagens dos diferentes bancos de dados sobre democracia ver Verkuilen (2009).

<sup>12</sup> King (1995) argumenta que “o método mais comum e cientificamente produtivo de construir a partir de pesquisa é replicar uma descoberta existente – seguir exatamente o caminho tomado por um

## Os resultados

Antes de proceder à análise dos dados, é importante apresentar algumas informações sobre a amostra. O banco de dados utilizado tem 12.078 entradas, contemplando o período entre 1946 e 2006. A unidade de análise são os países ao longo do tempo. Como nosso propósito é replicar os passos percorridos por Coppedge, Alvarez e Maldonado (2008), analisamos os dados de forma transversal durante o ano de 1985. É importante conferir a estatística descritiva das variáveis (observar possíveis *outliers*, erros de digitação, etc). Como pode ser observado, a amostra apresenta 127 casos, sugerindo que o critério mínimo de observações foi respeitado. O mesmo pode ser dito a respeito da proporção do número de casos por variável, 127 casos/10 variáveis, valor acima do mínimo recomendado. A Tabela 5 apresenta a matriz de correlação.

**Tabela 5**  
**Estatística descritiva**

<b>Variável</b>	<b>Média</b>	<b>Desvio padrão</b>	<b>N</b>	<b>Fonte</b>
V <sub>1</sub> (Liberdades Cívicas)	4,45	2,03	127	Freedom House
V <sub>2</sub> (Direitos Políticos)	4,30	2,25	127	Freedom House
V <sub>3</sub> (Competição)	22,01	24,82	127	Vanhanen (2007)
V <sub>4</sub> (Limites ao Executivo)	3,60	2,35	127	Polity
V <sub>5</sub> (Competitividade do Recrutamento do Executivo)	2,45	1,11	127	Polity
V <sub>6</sub> (Sufrágio adulto)	77,37	39,79	127	Bollen et al
V <sub>7</sub> (Seleção Legislativa)	8,86	3,03	127	Banks
V <sub>8</sub> (Direitos Políticos das Mulheres)	1,61	0,67	127	CIRI
V <sub>9</sub> (Participação)	29,16	23,77	127	Vanhanen (2007)
V <sub>10</sub> (Abertura do Recrutamento do Executivo)	3,20	1,52	127	Polity

A maior parte das correlações supera 0,30 (Tabela 6). Logo, seguindo a recomendação de Hair et al (2006) os dados são adequados a utilização da análise fatorial (as correlações com valores abaixo de 0,30 estão sombreadas). O

---

pesquisador anterior e então melhorar os dados ou metodologia, de uma maneira ou de outra” (KING, 1995, p. 445).

pesquisador deve atentar para o padrão de correlação entre as suas variáveis de interesse. Caso uma variável seja estatisticamente independente das demais, o pesquisador pode cogitar excluí-la da análise. Por exemplo, observa-se que a variável V<sub>8</sub> (Direitos Políticos das Mulheres) apresenta cinco correlações abaixo do patamar mínimo exigido. Em uma perspectiva mais conservadora, o pesquisador poderia excluí-la e estimar novamente a análise fatorial. Isso porque como a análise fatorial depende do padrão de correlação entre as variáveis observadas, espera-se que variáveis estatisticamente independentes não contribuam para a construção de um fator comum. Portanto, o teste final para a inclusão ou exclusão da variável será o nível de associação entre a variável e o fator extraído, sinalizado pelo valor da comunalidade. O próximo passo é verificar os testes de adequação da amostra. A tabela abaixo sintetiza essas informações.

**Tabela 6 - Matriz de correlações**

	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>	V <sub>7</sub>	V <sub>8</sub>	V <sub>9</sub>	V <sub>10</sub>
V <sub>1</sub>	1,000									
V <sub>2</sub>	,932	1,000								
V <sub>3</sub>	-,866	-,891	1,000							
V <sub>4</sub>	-,844	-,881	,863	1,000						
V <sub>5</sub>	-,811	-,883	,844	,917	1,000					
V <sub>6</sub>	-,353	-,398	,397	,373	,409	1,000				
V <sub>7</sub>	<b>-,251</b>	-,345	,336	,347	,382	,676	1,000			
V <sub>8</sub>	<b>-,185</b>	<b>-,180</b>	<b>,205</b>	<b>,242</b>	<b>,228</b>	,496	,422	1,000		
V <sub>9</sub>	-,491	-,477	,545	,512	,495	,629	,460	,420	1,000	
V <sub>10</sub>	-,338	-,422	,399	,524	,645	,440	,480	,312	,414	1,000

**Tabela 7 - Testes de adequação da amostra**

Teste	Valor observado
KMO	0,868
BTS	1254,06
gl	45
sig	0,000

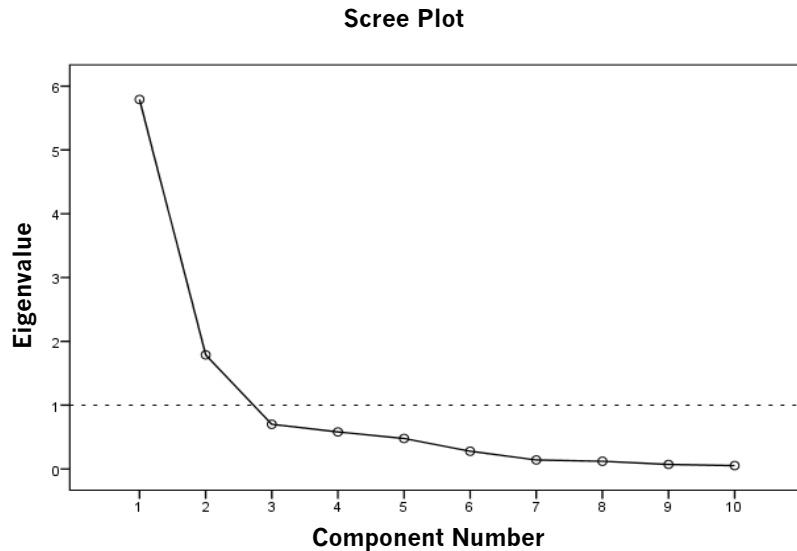
O valor do KMO da amostra é de 0,868, ou seja, superior ao patamar crítico de 0,60. Da mesma forma, o teste BTS é estatisticamente significativo ( $p < 0,000$ ). Em ambos os casos, os testes sugerem que os dados são adequados à análise fatorial. O passo seguinte é determinar o número de fatores que serão extraídos, a partir da Tabela 8.

**Tabela 8 - *Eigenvalues* e variância acumulada**

Componente	Initial eigenvalues			Extraction sums of squares loadings		
	Total	% variância	% acumulado	Total	% variância	% acumulado
1	5,791	57,906	57,906	5,791	57,906	57,906
2	1,789	17,885	75,791	1,789	17,885	75,791
3	0,700	6,996	82,788			
4	0,581	5,810	88,598			
5	0,478	4,776	93,374			
6	0,278	2,778	96,152			
7	0,140	1,404	97,556			
8	0,120	1,204	98,760			
9	0,071	,713	99,473			
10	0,053	,527	100,000			

O critério de Kaiser sugere que deve-se extrair dois fatores: o primeiro apresenta um *eigenvalue* de 5,791, carregando cerca de 58% da variância. O segundo fator apresenta *eigenvalue* de 1,789, carregando cerca de 18% da variância. Em conjunto, esses dois fatores explicam 75,79% da variância das variáveis originais. O gráfico 1 ilustra a dispersão dos componentes no *Scree test*.

**Gráfico 1**  
**Scree test**



A linha pontilhada ilustra o critério de Kaiser (*eigenvalue* >1). Ao utilizar o critério do *scree test* o pesquisador poderia ser levado a extrair mais fatores. Por exemplo, a extração de um terceiro fator apresentaria um ganho de 7% de variância. Tanto Hair et al (2006) quanto Schawb (2007) sugerem que a extração deve continuar até o pesquisador captar, pelo menos, 60% da variância. Por adotarmos uma perspectiva confirmatória, o critério da variância acumulada não deve ser utilizado com tanta rigidez já que definimos a priori o número de fatores que seriam extraídos. A Tabela 9 apresenta as comunalidades associadas a cada variável.



**Tabela 9**  
**Comunalidades**

Variável	Initial	Extraction
V <sub>1</sub> (Liberdades Civis)	1,000	0,880
V <sub>2</sub> (Direitos Políticos)	1,000	0,929
V <sub>3</sub> (Competição)	1,000	0,881
V <sub>4</sub> (Limitações ao Executivo)	1,000	0,902
V <sub>5</sub> (Competitividade do Recrutamento Executivo)	1,000	0,897
V <sub>6</sub> (Sufrágio Adulto)	1,000	0,757
V <sub>7</sub> (Seleção Legislativa)	1,000	0,675
V <sub>8</sub> (Direitos Políticos das Mulheres)	1,000	0,546
V <sub>9</sub> (Participação)	1,000	0,598
V <sub>10</sub> (Abertura do Recrutamento do Executivo)	1,000	0,473

De acordo com Schawb (2007), as comunalidades representam a proporção da variância para cada variável incluída na análise que é explicada pelos componentes extraídos. Por exemplo, os dois fatores extraídos explicam 88,0% da variância da V<sub>1</sub> (Liberdades Civis). Usualmente o valor mínimo aceitável é de 0,50. Logo, caso o pesquisador encontre alguma comunalidade abaixo desse patamar a variável deve ser excluída e a análise fatorial deve ser realizada novamente<sup>13</sup>. Além disso, baixa comunalidade entre um grupo de variáveis é um indício de que elas não estão linearmente correlacionadas e, por isso, não devem ser incluídas na análise fatorial. No caso acima, a variável V<sub>10</sub> (Abertura do Recrutamento Executivo) apresenta comunalidade ligeiramente abaixo do valor crítico. Em uma perspectiva mais conservadora, seria aconselhável excluí-la da análise e realizar novamente a análise fatorial. Para os nossos propósitos, vamos conservá-la.

Depois de eliminar as variáveis problemáticas (associadas ao baixo grau de comunalidade), o pesquisador deve analisar as cargas fatoriais de cada variável em relação aos componentes extraídos. A Tabela 10 sintetiza essas informações.

---

<sup>13</sup> Um procedimento adicional que o pesquisador pode adotar é analisar a matriz anti-imagem e verificar na diagonal os valores abaixo de 0,50. Escores abaixo desse patamar devem ser removidos da amostra e a análise fatorial deve ser realizada novamente (SCHAWB, 2007).

**Tabela 10**  
**Matriz não-rotacionada**

Variável	Componente	
	1	2
V <sub>1</sub> (Liberdades Cívicas)	-0,859	0,376
V <sub>2</sub> (Direitos Políticos)	-0,902	0,339
V <sub>3</sub> (Competição)	0,890	-0,299
V <sub>4</sub> (Limitações ao Executivo)	0,908	-0,280
V <sub>5</sub> (Competitividade do Recrutamento Executivo)	0,918	-0,234
V <sub>6</sub> (Sufrágio Adulto)	<b>0,633</b>	<b>0,597</b>
V <sub>7</sub> (Seleção Legislativa)	<b>0,567</b>	<b>0,594</b>
V <sub>8</sub> (Direitos Políticos das Mulheres)	<b>0,413</b>	<b>0,613</b>
V <sub>9</sub> (Participação)	0,699	0,331
V <sub>10</sub> (Abertura do Recrutamento do Executivo)	0,638	0,257

Na matriz não-rotacionada, o pesquisador deve: (1) observar as cargas fatoriais de cada variável e (2) identificar as variáveis que apresentam elevadas cargas fatoriais em ambos os componentes, no caso, valores acima de 0,40. Tanto Hair et al (2006) quanto Schawb (2007) sugerem que um dos pressupostos da análise fatorial é a estrutura simples de seus componentes. O argumento defende que a mesma variável não pode contribuir para a construção de fatores distintos. Adota-se 0,40 como limite aceitável da contribuição da variável na criação do fator com o objetivo de evitar o problema da indeterminação da relação entre variáveis e fatores. Em uma perspectiva exploratória, critérios puramente técnicos podem ajudar o pesquisador a tomar essa decisão. Em uma perspectiva confirmatória, no entanto, é prudente o pesquisador utilizar elementos teóricos para justificar a inclusão e a exclusão de variáveis em sua análise. O passo seguinte é rotacionar as variáveis em relação aos componentes extraídos. Com o objetivo é facilitar a visualização da relação entre as variáveis observadas e os componentes extraídos. A Tabela 11 apresenta os valores de cada componente após a rotação.

**Tabela 11**  
**Matriz rotacionada (VARIMAX) (V<sub>1</sub> até V<sub>10</sub>)**

Variável	Componente	
	1	2
V <sub>1</sub> (Liberdades Cívicas)	-0,927	-0,141
V <sub>2</sub> (Direitos Políticos)	-0,944	-0,195
V <sub>3</sub> (Competição)	0,912	0,223
V <sub>4</sub> (Limitações ao Executivo)	0,917	0,248
V <sub>5</sub> (Competitividade do Recrutamento Executivo)	0,901	0,293
V <sub>6</sub> (Sufrágio Adulto)	0,216	0,843
V <sub>7</sub> (Seleção Legislativa)	0,162	0,806
V <sub>8</sub> (Direitos Políticos das Mulheres)	0,021	0,739
V <sub>9</sub> (Participação)	<b>0,414</b>	<b>0,653</b>
V <sub>10</sub> (Abertura do Recrutamento do Executivo)	<b>0,402</b>	<b>0,558</b>

No caso acima, duas variáveis violam ligeiramente esse pressuposto: V<sub>9</sub> (Participação) e V<sub>10</sub> (Abertura do Recrutamento Executivo). Elas devem ser excluídas e a análise fatorial deve ser realizada novamente. A Tabela 12 ilustra esses dados<sup>14</sup>.

---

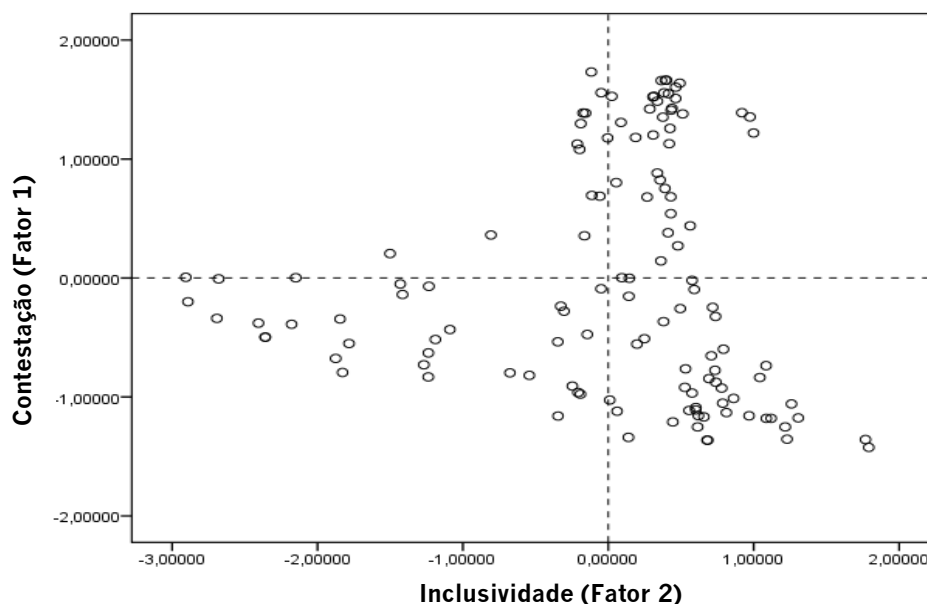
<sup>14</sup> As estatísticas de interesse referentes ao modelo 2 estão em anexo.

**Tabela 12**  
**Matriz rotacionada (VARIMAX) (V<sub>1</sub> até V<sub>8</sub>)**

<b>Variável</b>	<b>Componente</b>	
	<b>1</b>	<b>2</b>
V <sub>1</sub> (Liberdades Cívicas)	<b>-0,934</b>	-,120
V <sub>2</sub> (Direitos Políticos)	<b>-0,953</b>	-,176
V <sub>3</sub> (Competição)	<b>0,922</b>	,191
V <sub>4</sub> (Limitações ao Executivo)	<b>0,927</b>	,203
V <sub>5</sub> (Competitividade do Recrutamento Executivo)	<b>0,911</b>	,232
V <sub>6</sub> (Sufrágio Adulto)	,257	<b>0,840</b>
V <sub>7</sub> (Seleção Legislativa)	,200	<b>0,824</b>
V <sub>8</sub> (Direitos Políticos das Mulheres)	,056	<b>0,770</b>

Seguindo a Tabela 12, nenhuma variável apresenta estrutura complexa, ou seja, cargas fatoriais acima de 0,40 em ambos os componentes; e que houve um incremento generalizado no valor das cargas fatoriais associadas aos componentes extraídos. Além disso, ao se comparar os resultados encontrados utilizando todas as variáveis e os resultados obtidos utilizando as oito variáveis, chega-se à conclusão de que a segunda solução além de ser mais parcimoniosa, apresenta fatores que carregam maior quantidade de variância acumulada: 75,79% no primeiro modelo versus 82,30% no segundo modelo. Em termos substantivos, o primeiro fator está relacionado às variáveis V<sub>1</sub>, V<sub>2</sub>, V<sub>3</sub>, V<sub>4</sub> e V<sub>5</sub>, representando a dimensão da contestação. O segundo fator está associado às variáveis V<sub>6</sub>, V<sub>7</sub> e V<sub>8</sub>, ilustrando a dimensão da inclusão. O Gráfico 2 ilustra a distribuição dos países a partir da interação entre essas duas dimensões.

**Gráfico 2**  
**Contestação e Inclusividade**



Os valores são padronizados de tal forma que a média é zero e a distância entre os escores é medida em termos de desvio padrão. No quadrante inferior-esquerdo, tanto a contestação (fator 1) quanto a inclusividade (fator 2) estão abaixo da média, sugerindo que esses regimes se aproximam do que Dahl (1971) denominou de hegemonia fechada. No quadrante inferior-direito, a inclusão está acima da média, mas a contestação é fracamente desenvolvida o que caracteriza os regimes como hegemonias inclusivas. No quadrante superior-esquerdo estão os países com grau de contestação acima da média, mas que apresentam menor nível de inclusividade, casos que Dahl (1971) define como oligarquias competitivas. Finalmente, o quadrante superior-direito apresenta os países com ambas as dimensões, inclusividade e contestação, acima da média, ou seja, regimes poliárquicos.

### **Conclusão**

Como mensurar fenômenos que não podem ser diretamente observados? Essa foi a questão de pesquisa que orientou nossos esforços. Partindo do pressuposto de que a operacionalização de conceitos em variáveis empiricamente observáveis é um procedimento central para viabilizar a produção do conhecimento,

esse artigo apresentou a lógica da análise fatorial a partir de uma abordagem intuitiva. Isso porque grande parte da literatura apresenta um grau proibitivo de complexidade, o que dificulta a compreensão. Em termos substantivos, acreditamos que problemas de mensuração constituem um dos principais obstáculos ao desenvolvimento do conhecimento. De forma direta, acreditamos que variáveis mal medidas comprometem a validade das inferências. Quanto mais erros em nossas medidas, tanto piores serão nossas conclusões a respeito dos fenômenos de interesse.

Acreditamos fortemente que a credibilidade dos métodos empregados é um critério central para avaliar os resultados de pesquisa (COLLIER, SEAWRIGHT e MUNCK, 2004, p.23), pois é exatamente a utilização rigorosa do método que distingue o conhecimento sistemático de outras formas de conhecimento. Esperamos contribuir para a difusão da análise fatorial como técnica de mensuração nas Ciências Sociais em geral e na Ciência Política em particular. Afinal, medidas confiáveis e válidas constituem um elemento fundamental na construção do conhecimento científico.

### **Referências Bibliográficas**

ASHER, H. B. *Causal Modeling*. Beverly Hills, CA: Sage, 1983.

BAKER, E. K. Polyarchy Plus: Measuring the Quality of Democracy within Thirteen Eastern European Countries. Trabalho apresentado na *American Political Science Association*, São Francisco, CA, 30 de Agosto - 2 de setembro, 2001.

BARTHOLOMEW, D.J. The foundations of factor analysis, *Biometrika*, 71, 221-232, 1984.

BLALOCK, H. M. *Measurement in the social sciences: Theories and strategies*. Chicago, Illinois: Aldine Publishing Company, 1974.

\_\_\_\_\_. *Basic Dilemmas in the Social Sciences*. Beverly Hills, CA: Sage, 1984.

BOLLEN, K. A.; ARMINGER, G. Observational Residuals in Factor Analysis and Structural Equation Models. *Sociological Methodology*, 21, 235-262, 1991.

BOLLEN, K. A. *Structural Equations with Latent Variables*. Wiley Series in Probability and Mathematical Statistics. Nova York: Wiley, 1989.

BOLLEN, K.A.; CURRAN, P. J. *Latent Curve Models: A Structural Equation Perspective*. Wiley Series in Probability and Mathematical Statistics. Nova York: Wiley, 2006.

BOLLEN, K.A.; LONG, J.S. *Testing Structural Equation Models*. Newbury Park, CA: Sage, 1993.

BOLLEN, K. A.; GRANDJEAN, B. D. "The Dimension(s) of Democracy: Further Issues in the Measurement and Effects of Political Democracy." *American Sociological Review*, 46, 5, 651-59, 1981.

BONJEAN, C. M.; BROWNING, H. L. Toward Comparative Community Research: A Factor Analysis of United States Counties. *The Sociological Quarterly*, 10, 2, 157-176, 1969.

\_\_\_\_\_. "The Scree Test for the Number of Factors." *Multivariate Behavioral Research*, 1(2), 245-276, 1966.

CATTELL, R. B. *The Scientific Use of Factor Analysis in Behavioral and Life Sciences*. Nova York: Plenum, 1978.

COLLIER, D; SEAWRIGHT, J, MUNCK; GERARDO, L. Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology. In: BRADY, H. e COLLIER, D. (orgs), *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowman and Littlefield, 2004.

COOPER, J. C.B. *Factor Analysis: An Overview*. *The American Statistician*, 37, 2, 141-147, 1983.

COPPEDGE, M. "Two Persistent Dimensions of Democracy: Contestation and Inclusiveness." *Journal of Politics*, 70, 3, 1-45, 2008.

COSTELLO, A. B; OSBORNE, J. W. "Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis." *Practical Assessment Research & Evaluation*, 10, 7, 13-24, 2005.

DAHL, R. *Poliarquia: Participação e Oposição*. São Paulo: Edusp, 1971.

DANCEY, C; REIDY, J. *Estatística Sem Matemática para Psicologia: Usando SPSS para Windows*. Porto Alegre: Artmed, 2006.

DECOSTER, J. *Overview of Factor Analysis*. [Online] Disponível em: <<http://www.stat-help.com/notes.html>> Acesso em: [22 jan. 2010].

DUNN, M. J.; SCHNECK, R.; LAWSON, J. "A Test of the Uni-Dimensionality of Various Political Scales through Factor Analysis: A Research Note." *Canadian Journal of Political Science / Revue Canadienne de Science Politique*, 6, 4, 664-669, 1973.

DUNTEMAN, G. H. *Principal Components Analysis*. Newbury Park: Sage, 1989.

FIELD, A. *Discovering Statistics Using SPSS*. Londres: Sage, 2005.

GARSON, G. D. *Statnotes: Topics in Multivariate Analysis*. [Online] Disponível em: <<http://faculty.chass.ncsu.edu/garson/PA765/statnote.htm>> Acesso em: [22 jan. 2010].

GRUMM, J. G. "A Factor Analysis of Legislative Behavior." *Midwest Journal of Political Science*, 7, 4, 336-356, 1963.

HAIR, Jr; BLACK, W. C; BABIN, B. J; ANDERSON, R. E e TATHAM, R. L. *Multivariate Data Analysis*. 6ª edição. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.

HARMAN, H.H. *Modern Factor Analysis*. 2ª edição. Chicago: University of Chicago Press, 1967.

HO, R. *Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS*. North Carolina: Chapman and Hall, 2006.

ISOGAWA, Y; OKAMOTO, M. "Linear Prediction in the Factor Analysis Model." *Biometrika*, 67, 2, 482-484, 1980.

KAPLUNOVSKY, A.S. Why using factor analysis? (dedicated to the centenary of factor analysis). [Online] Disponível em: <<http://www.magniel.com/fa/kaplunovsky.pdf>> Acesso em: [22 jan. 2010].

KIM, J; MUELLER, C. W. *Factor analysis: Statistical methods and practical issues*. Beverly Hills, CA: Sage, 1978a.

\_\_\_\_\_. *Introduction to factor analysis - what it is and how to do it*. Beverly Hills, CA: Sage, 1978b.

KING, G. How not to lie with statistics [Online] Disponível em: <<http://gking.harvard.edu/files/mist.pdf>> Acesso em: [22 jan. 2010].

KLINE, R. B. *Principles and Practice of Structural Equation Modeling*. Nova York: Guilford, 2004.

LAWLEY, D. N; MAXWELL, A. E. "Regression and Factor Analysis." *Biometrika*, 60, 2, 331-338, 1973.

LIGNY, C. L; NIEUWDORP, G. H. E; BREDERODE, W. K; HAMMERS, W. E; HOUWELINGEN, J. C. van. An Application of Factor Analysis with Missing Data. *Technometrics*, 23, 1, 91-95, 1981.

MACKELPRANG, A. J. "Missing Data in Factor Analysis and Multiple Regression." *Midwest Journal of Political Science*, 14, 3, 493-505, 1970.



- MOONEY, C. Z. "Bootstrap Statistical Inference: Examples and Evaluation For Political Science." *American Journal of Political Science*, 40, 570-602, 1996.
- PALLANT, J. *SPSS Survival Manual*. Open University Press, 2007.
- ROBERTS, B. R. "A Confirmatory Factor-Analytic Model of Alienation." *Social Psychology Quarterly*, 50, 4, 346-351, 1987.
- RUMMEL, R. J. "Understanding Factor Analysis." *The Journal of Conflict Resolution*, 11, 4, 444-480, 1967.
- RUMMEL, R.J. *Applied Factor Analysis*. Evanston: Northwestern University Press, 1970.
- SANTOS, M. H; COUTINHO, M. "Política comparada: estado das artes e perspectivas no Brasil." *BIB*, 5 4, 3-146, 2000.
- SCHAWB, A.J. *Eletronic Classroom*. [Online] Disponível em: <<http://www.utexas.edu/ssw/eclassroom/schwab.html>> Acesso em: [22 jan. 2010].
- SLATIN, G. T. "A Factor Analytic Comparison of Ecological and Individual Correlations: Some Methodological Implications." *The Sociological Quarterly*, 15, 4, 507-520, 1974.
- SOARES, G. "O calcanhar metodológico da ciência política no Brasil." *Sociologia*, 48, 27-52, 2005.
- SPEARMAN, C. General intelligence, objectively determined and measured. *American Journal of Psychology*, 15, 201-293, 1904.
- TABACHNICK, B.; FIDELL, L. *Using multivariate analysis*. Needham Heights: Allyn & Bacon, 2007.
- THURSTONE, L. L. *The vector of mind*. Chicago: University of Chicago, 1935.
- VALLE E SILVA, N. *Relatório de Consultoria sobre Melhoria do Treinamento em Ciência Social Quantitativa e Aplicada no Brasil*. Rio de Janeiro, Laboratório Nacional de Computação Científica, 1999.
- VERMUNT, J. K; MAGIDSON, J. Factor Analysis with categorical indicators: A comparison between traditional and latent class approaches. In: VAN DER ARK, A. CROON, M.A. and SIJTSMA, K. *New Developments in Categorical Data Analysis for the Social and Behavioral Sciences*. Mahwah: Erlbaum, 2005.

WERNECK VIANNA, L. et al. "Doutores e teses em ciências sociais." *Dados*, 41, 3, 453-515, 1998.

YALCIN, I; AMEMIYA, Y. Nonlinear Factor Analysis as a Statistical Method. *Statistical Science*, 16, 3, 275-294, 2001.

ZELLER, R. A; CARMINES, E. G. *Measurement in the social sciences: The link between theory and data*. Cambridge: Cambridge University Press, 1980.

---

Dalson Brito Figueiredo Filho - dalsonbritto@yahoo.com.br

José Alexandre da Silva Júnior - jasjunior2007@yahoo.com.br

Recebido para publicação em fevereiro de 2009  
Aprovados para publicação em abril de 2010