

# **Numa Escala De 1 A 10, Quanto Dói?**

## **Sobre A Construção De Índices E Sua Validade Nas Ciências Sociais**

Haína Coelho. Doutoranda em Ciência Política Pela UFPE. E-mail: haina\_sama@hotmail.com

### **Resumo**

Este artigo fornece um guia metodológico para a elaboração de índices. Esclareço sobre os tipos de índices de acordo com o nível de mensuração das variáveis e faço uma discussão sobre a aplicação de indicadores compostos nas ciências sociais utilizando o debate de normatividade na ciência. São dadas noções iniciais sobre teoria de mensuração, e sua utilização na construção de índices. Cada passo para criar um índice é detalhado: desde seleção de itens, passando por análise de relações empíricas, ponderação dos itens e validação do indicador, acompanhados eventualmente de códigos de execução das técnicas nos softwares SPSS, Stata e R. Para finalizar, são levantadas alguns obstáculos que podem aparecer no caminho para a construção de índices nas ciências sociais, como dados inexistentes.

**Palavras-Chave:** Índices; Validade; Mensuração; Ciências Sociais.

### **Abstract**

The goal of this paper is to provide a methodological guide for elaboration of indexes, or composite indicators. I clarify about types of indexes according to their data measurement, and I discuss about its use in the social sciences employing the debate about normativity on science. I give introductory notions on measurement theory, and its use on indexes elaboration. Each step to build a valid index is detailed: from item selection, going through empirical relation analysis, item scoring and finally validation of the index, eventually followed by execution codes for the techniques in softwares SPSS, Stata and R. To conclude, I raise some extra questions that could appear on the way to build an index in the social sciences, such as missing data.

**Key-words:** Indexes; Validity; Measurement; Social Sciences.

# Numa Escala De 1 A 10, Quanto Dói? Sobre A Construção De Índices E Sua Validade Nas Ciências Sociais

Haína Coelho – Universidade Federal de Pernambuco

*"A major difference between a "well-developed" science such as physics and some of the less "well-developed" sciences such as psychology or sociology is the degree to which things are measured."*

*(Fred Roberts)*

## 1. INTRODUÇÃO

Em emergências hospitalares, para elencar a prioridade de atendimento, o paciente passa por uma triagem onde é feita sua classificação de risco. Esta é elaborada com base em alguns indicadores. Por exemplo, existe um índice que é internacionalmente utilizado para avaliação de risco em hospitais, o Índice de Severidade de Emergência. Ele considera se o paciente precisa de intervenção imediata, se ele está sentindo muita dor ou corre risco de piorar, e se recursos adicionais à consulta - como exames laboratoriais - serão necessários (*ESI Triage Research Team*, 2004). Há outros índices usados pela medicina com este mesmo fim.

Como o leitor atento pode deduzir, apesar de os dados serem produzidos primariamente pela triagem, eles são puramente observacionais. Afinal de contas, é impossível estimar exatamente o nível de dor que outra pessoa está sentindo, por mais que ela própria a descreva. Isso pode levar a erros de mensuração por parte da enfermagem, classificando um paciente de forma errada na escala de risco. Os índices nas ciências sociais funcionam de forma semelhante.

Índices são medidas ordinais compostas (BABBIE, 2010), ou seja, mensuram um conceito em gradação de níveis a partir dos escores de duas ou mais variáveis que componham este conceito. Os escores podem tanto ser contínuos como dicotômicos. Os escores de todas as variáveis são agregados em uma única medida, o índice.

Alguns desses índices são muito famosos. O IDH (índice de Desenvolvimento Humano) é um deles, e calcula níveis de desenvolvimento humano de países através de quatro indicadores: taxa de alfabetização, taxa de escolarização, expectativa de vida e PIB *per capita*. Em caminho similar, há estudos buscando agregar para análise indicadores altamente disseminados de qualidade de vida, como esperança de vida e nutrição da população (SANTOS, TEODORO & FACCIN, 2015).

Índices que mensuram níveis de educação, como o IDEB (Índice de Desenvolvimento da Educação Básica), são úteis como referência a governantes e gestores públicos na melhoria de políticas públicas.

Um provável obstáculo à criação de índices nas ciências sociais está na complexidade e fluidez dos conceitos, e principalmente na dificuldade de mensuração. Teoria de mensuração é uma necessidade ao desenvolvimento empírico de qualquer disciplina, e onde está um dos grandes calos das ciências sociais.

Este trabalho tem como objetivo prover um guia introdutório à construção de índices e noções básicas de teoria de mensuração. O público-alvo é composto de estudantes de graduação e pós-graduação de todas as áreas, mas com direcionamento maior aos das Ciências Sociais, em especial Ciência Política. Um iniciante em metodologia científica poderá usar este artigo como referencial básico para construção de índices, e nele será direcionado a fontes que aprofundem suas necessidades específicas.

Difundir de forma sistemática os conhecimentos necessários para aprimoramento da transformação de conceitos teóricos em medidas tangíveis e testáveis é a melhor forma de promover o desenvolvimento empírico de nossa área de conhecimento. O artigo, tentando alcançar os indivíduos em estágio inicial, é uma missão neste sentido.

Na seção a seguir, será esclarecida a natureza de um índice e seus usos. A seção subsequente entrará brevemente no debate da validade dessa técnica nas ciências sociais, e, por fim, trago as etapas para a construção adequada de um índice, antes das considerações finais.

## **2. ÍNDICES: O QUE SÃO, PARA QUE SERVEM E COMO SÃO CALCULADOS?**

Para entender o que são índices sigo um passo-a-passo lógico, iniciando pelo princípio, por isso peço ao leitor entediado um pouco de paciência. A teoria é quase sempre o ponto de partida nas pesquisas científicas, e é de seu desenvolvimento que extraímos uma série de conceitos abstratos (VAN EVERA, 1997). Quando valores são atribuídos a esses conceitos, eles se tornam as variáveis da pesquisa (VAN EVERA, 1997).

Contudo, a parte interessante da ciência é comparar variáveis e estabelecer conexões entre as mesmas, especialmente as de cunho causal. Para que isso seja possível, entramos com as relações<sup>1</sup> matemáticas (MOORE & SIEGEL, 2013).

---

<sup>1</sup> No inglês é feita uma diferenciação entre *relationship* e *relation*, em que esta última se refere ao termo matemático. Uma vez que ambos são traduzidos como *relação* neste contexto, o *paper* usa *relação* apenas para o termo matemático, utilizando-se de variantes semânticas quando não for o caso.

Uma relação entre conceitos existe quando a partir de uma variável podemos obter informações sobre outra variável, e assim melhor descrever o universo que estamos observando (MOORE & SIEGEL, 2013). É importante, então, que saibamos representar com acuidade essas relações, ou estaríamos executando uma má explicação da realidade.

Mensuração é a etapa crucial para isso, pois por ela atribuímos os valores mais adequados para preservar as relações entre variáveis - o que na matemática se transforma em uma função manipulável (ROBERTS, 2009). A mensuração representa quantidades, ou define em quais categorias os atributos de certos objetos são enquadrados (NUNNALLY & BERNSTEIN, 1978). As duas variações respectivamente vinculados à mensuração de atributos diretamente mensuráveis, como altura, e categorias abstratas, como inteligência (NUNNALLY & BERNSTEIN, 1978).

Qualquer cientista desenvolve medidas que obedecem a certas regras para quantificar atributos, e se utilizam de sistemas matemáticos para analisar a estrutura dos dados. (NUNNALLY & BERNSTEIN, 1978). Isso se diferencia da mensuração puramente matemática pela conexão com o universo empírico: um sistema matemático só necessita de validação interna de regras para fazer sentido (NUNNALLY & BERNSTEIN, 1978).

O processo de mensuração, tratamento matemático dos conceitos, é o que permite à Ciência Política que seus conceitos sejam transformados em variáveis propriamente ditas, e não somente sejam alocados em classificações categóricas (SARTORI, 1970). Os atributos pertencentes aos conceitos devem ser aptos de gradação, o que é diferente de ingenuamente atribuir valores numéricos a categorias (SARTORI, 1970).

Não há uma única forma de mensurar conceitos, pois isso depende do tipo de dado que a pesquisa está lidando. De forma simplista, caso os dados tenha valores categóricos, eles são chamados de discretos. Já quando os valores são números reais, eles são dados contínuos. Para dados discretos, a mensuração pode ser nominal, quando é puramente categórica, ou ordinal, em que há hierarquia entre as categorias. Funcionando para os dois tipos de dado temos a mensuração intervalar, em que o espaço entre valores é sempre o mesmo, e a de razão, executada da mesma forma que a intervalar, mas considerando o ponto zero como indicador de ausência de efeito (Gill, 2006). O Quadro 1 sumariza essas informações.

### Quadro 1 - Tipos de índice

Nível de mensuração	Tipo de dado	Exemplo
Nominal	categórico	Guerra/paz
Ordinal	categórico	Posicionamento na escala ideológica
Intervalar	categórico ou contínuo	Percepção do eleitor sobre a corrupção
Razão	categórico ou contínuo	Número de vetos pelo Executivo

**Fonte:** elaboração própria, compilado com base em Gill (2006), Moore e Siegel (2013), Roberts (2009).

Nas ciências sociais, destacando a Ciência Política, o processo de mensuração é mais complexo devido ao caráter multidimensional de muitos conceitos. Por exemplo, o conceito de participação política, cujos atributos abrem a possibilidade de existência de diversas modalidades de participação (HANSEN, 2016).

Este aspecto leva os pesquisadores a identificarem indicadores para esses conceitos, adotando as mais diversas estratégias para medi-los, seja de forma dicotômica, policotômica, por intervalos escalares, etc. (LANDMAN, 2008).

Os índices, então, surgem advindos de mensuração multivariada desses indicadores (HAIR *et al*, 2013). É importante diferenciar índices e escalas, apesar deste fator não ser tão relevante para este artigo. Uma escala se utiliza das diferenças de intensidade nos indicadores de uma variável para identificar padrões, enquanto que um índice acumula os valores de cada indicador para ranquear as observações de forma mais geral (BABBIE, 2010).

A utilização de índices e escalas é recorrente na Ciência Política. Nos estudos de democracia isso é bastante evidente, como exemplo os trabalhos de classificação de regimes (MAINWARING, BRINKS & PEREZ-LIÑAN, 2001; DAHL, 1997) qualidade de democracia (DEMOCRACY BAROMETER, 2016; FREEDOM HOUSE, 2016), e podemos mencionar também estudos sobre desenhos institucionais, como independência do banco central (LIJPHART, 2008; CUKIERMAN, WEBB & NEYAPTI, 1992), rigidez constitucional (LORENZ, 2005; LIJPHART, 2008), e governança eleitoral (MOZZAFFAR & SCHEDLER, 2002; TAROUCO, 2012).

O primeiro ponto destacado no artigo foi a dificuldade que enfermeiros possuem para mensurar um dado que não pode ser diretamente observado, e que isso se repete nas ciências sociais. Gill (2006) alerta como é frequente a construção de índices artificiais por esses pesquisadores para medir algumas características da sociedade através de dados ordinais. Qual a real natureza deste problema? A próxima seção examina essa questão.

### 3. PODEMOS USAR ÍNDICES?

Como foi dito na seção anterior, a matemática transforma em funções as relações existentes entre variáveis para que seja possível manipulá-las e assim compará-las. Entretanto, alguns dados não podem ser manipulados, como os sintomas de um paciente na emergência, o desenho institucional de um país, as características de um regime democrático, ou a implementação de uma política pública. Essa é a realidade para muitas informações consideradas relevantes nas ciências sociais.

Essa é uma das grandes questões metodológicas apresentadas por Brady e Collier (2010): a dicotomia entre dados observacionais e dados experimentais. Ao trabalhar com dados observacionais, não ter a habilidade de manipulá-los significa não poder controlar seu efeito sobre a relação analisada. Dados observacionais muitas vezes também não são objetivos de se mensurar.

Este fator levanta um dos grandes *trade-offs* de uma pesquisa científica: confiabilidade *versus* validade. Confiabilidade existe quando diferentes mensurações do mesmo fenômeno geram os mesmos resultados, e validade quando as mensurações refletem o que o pesquisador quer medir (KING, KEOHANE & VERBA, 1994). A partir de agora uso democracia como exemplo, devido a sua icônica aplicabilidade. No entanto, o conceito poderia ser X, Y ou Z, portanto se algum outro é mais lógico para você, faça o exercício mental de substituir os termos.

Pensando assim para o caso de democracia, se escolhermos um conceito policotômico e quisermos comparar entre países, ou uma série temporal de um único país, temos que escolher indicadores que acusem certo nível democrático em uma observação todas as vezes que mensurarmos essa observação, e esses indicadores devem de fato apontar o nível democrático.

É a controvérsia a que se referem Kellstedt e Whitten (2009) quando afirmam que por mais que seja tentador pensar em democracia como uma gravidez (é ou não é), ela funciona como um *continuum* de níveis, e a parte difícil é identificá-los. Além de o *tradeoff* naturalmente existir, o conceito multidimensional de democracia, e a não manipulabilidade de seus indicadores tornam essa tarefa bastante complexa.

Neste ponto o leitor se pergunta se vale a pena tal esforço, ou se os índices construídos significarão alguma coisa cientificamente. Para responder a primeira reflexão, vamos retomar o debate primário sobre a normatividade na Ciência Política, e pedir que o leitor acompanhe mais uma digressão.

Weber (1982), de certo modo, se fez a mesma pergunta. Ele não se referia a índices em específico, obviamente, mas da compreensão e interpretação de fenômenos políticos de forma geral.

A ciência não responde como esses objetos são compensadores o suficiente para o esforço de conhecê-los. O valor de que está embebida uma posição política não deve jamais ser levada para dentro de sala, pois corre-se o risco de transformar a ciência em demagogia (WEBER, 1982 [1946]).

Este argumento em nada ajuda a primeira reflexão do leitor, quiçá a reforça. Mas em diálogo essas ideias weberianas, Leo Strauss (2009 [1953]) traz uma nova interpretação. Ele aponta como a posição de diferenciação de fatos e valores não é uma rejeição ao dever-ser, mas a rejeição de que é possível conhecer o dever-ser, que para explicar um fenômeno social devemos conhecê-lo pelo que ele é. O que Strauss alega, entretanto, é que é difícil compreender um fenômeno social sem um parâmetro conceitual prévio. Quando rejeitamos juízos de valor, estamos evitando “chamar as coisas pelo seu nome” como ele coloca, e ameaçamos inclusive a objetividade da interpretação. Os objetos das ciências sociais, para fechar seu debate com Weber, surgem através da referência de valores. Se retomarmos nosso exemplo de qualidade de democracia, isso fica mais evidente.

O que Strauss está sugerindo é o que motiva mais tarde o debate analítico e epistemológico da democracia pragmática: a dicotomização entre normatividade e empiria pela Ciência Política contemporânea deve ser superada. É possível endossar valores sem abrir mão dos fatos, e é útil para a disciplina, devido sua natureza, "conciliar a necessidade empírica de explicar as instituições políticas com a ambição normativa de propor formas de organização política mais desejáveis" (POGREBINSCHI, 2010, p. 689).

Ou seja, a compensação que Weber questiona para o estudo de fenômenos sociais está na normatividade: não no dever-ser do objeto, mas na escolha do mesmo - por isso ainda estudamos qualidade de democracia, a despeito do caminho espinhoso e das severas críticas que serão tomadas nesse respeito. O uso de índices acompanha este resultado lógico devido ao já mencionado caráter multidimensional dos conceitos que o tornam necessário.

Embora não possamos medir a eficiência democrática de um desenho institucional tal qual o desempenho de um carro, é importante diferenciar o que funciona ou não. Isso nos leva à segunda reflexão do leitor: essas análises têm alguma utilidade? A isso respondo: podem ter, se seguirem uma série de critérios rigorosos. A próxima seção desenvolve esses critérios de forma a tentar superar os problemas com os índices nas ciências sociais.

#### **4. COMO CRIAR UM ÍNDICE VÁLIDO?**

Embora diversos pesquisadores tenham trabalhado este tema, ainda que de forma dispersa, vou aqui sistematizar o processo de construção de índices, reproduzindo as quatro etapas descritas

por Babbie (2010) - seleção de itens, análise das relações empíricas, pontuação do índice, e sua validação - e aprofundando a discussão metodológica em cada uma delas.

#### 4.1 Seleção De Itens

É importante esclarecer que seleção de itens se refere aos componentes do índice, não aos casos da pesquisa. A primeira etapa está diretamente ligada ao uso de conceitos teoricamente orientados. Para que o índice represente o conceito que desejamos, é preciso escolher variáveis que preservem o sentido do conceito. Para isso, o conceito deve estar bem delimitado. Conceitos muito abstratos são de pouca utilidade em índices uma vez que dificultam a operacionalização das variáveis.

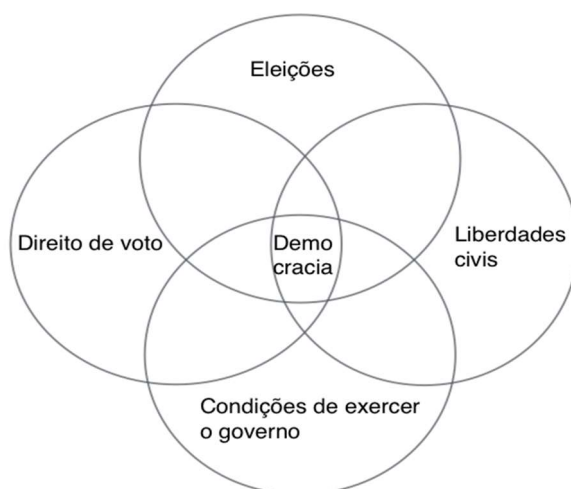
Por exemplo, como mensurar o conceito de alienação? Aliás, o que é alienação? Um dos problemas de conceitos na Ciência Política é que muitas vezes eles denotam amplitude o suficiente para prejudicar sua precisão empírica, o chamado estiramento conceitual, ou *conceptual stretching* (Sartori, 1970).

Estiramento conceitual é diferente da generalização de um conceito pela diminuição de atributos, de forma que ele seja mais inclusivo: se trata de deixar o conceito amplo mas sem muitas especificidades. Ou seja, vago e obscuro (Sartori, 1970).

É preciso que o conceito esteja bem definido e que se escolha se os indicadores se referem a ele de forma geral ou a suas dimensões de forma específica (Babbie, 2010). Uma forma de trabalhar essas decisões cruciais com clareza é utilizar diagramas de Venn. Para o conceito de forma geral, selecionar as condições necessárias e suficientes que o formam (Mahoney e Vanderpoel, 2014). Essas condições já podem ser interpretadas como os indicadores, se operacionalizado dessa forma, e visualmente os diagramas caracterizam uma relação de interseção de conjuntos, como na Figura 1 abaixo:



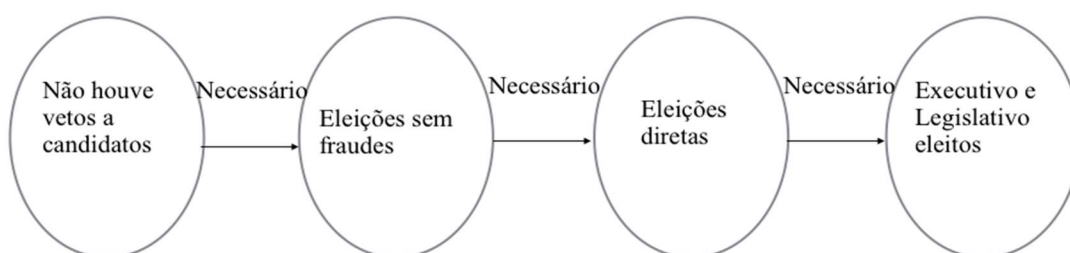
**Figura 1 - Condições Necessárias E Suficientes Para A Democracia**



Fonte: Elaboração própria, com base em Mainwaring, Brinks, Perez-Liñan (2001).

Para o caso de uma dimensão específica, o método de elaboração em sequência funciona bem. Ele organiza os conjuntos em uma cadeia de condições necessárias ou suficientes (Mahoney e Vanderpoel, 2014). A lógica é que para cada dimensão possa existir uma gradação. Continuando com o exemplo de democracia, posso escolher uma dimensão na Figura 2:

**Figura 2 - Condições necessárias para a dimensão Eleições**



Fonte: Elaboração própria, com base em Mainwaring, Brinks e Perez-Liñan (2001).

Ao garantir que os indicadores representem uma única dimensão cada, evita-se que conceitos mais gerais tenham dimensões repetidas entre si. Essa tática também pode funcionar como uma forma de separar os componentes endógenos de exógenos, para qual alerta King, Keohane e Verba (1994). É bom também garantir variância dentro dos indicadores, pois se um deles esteja construído

de tal forma que todos os casos ou nenhum caso seja incluído, ele não terá utilidade para o índice (BABBIE, 2010). Um exemplo empírico de alto refinamento de seleção de itens para construção de índice é a árvore conceitual do *Democracy Barometer* (MERKEL, BOCHSLER, *et al*, 2016).

Uma estratégia de seleção de itens puramente estatística é a análise de agrupamento, ou de *clusters* como é mais conhecida. Essa técnica identifica quais variáveis não alteram os resultados caso sejam retiradas da análise através de medidas de similaridade entre os objetos (Kubrusly, 2001).

Para uma análise básica de agrupamento, os seguintes códigos de *software* podem ser usados:

#### **Código no SPSS:**

```
Menu Analyze > Classify > K-Means Cluster/Hierarchical Cluster/Two-step Cluster
```

#### **Código no Stata:**

```
cluster complete v1 v2 v3, measure(absolute)
```

#### **Código no R:**

```
d <- dist(dados, method = "euclidean")  
fit <- hclust(d, method="ward")
```

Lembro que esses códigos podem receber diversas especificações a depender das necessidades de análise do pesquisador.

## **4.2 Análise De Relações Empíricas**

Escolhidos os indicadores com base nos conceitos teóricos, o próximo passo é verificar se esses indicadores apresentam alguma relação entre si, seja ela bivariada ou multivariada (Babbie, 2010). A diferença desta etapa para a anterior é que ao selecionar os itens estamos preocupados com sua relevância individual, e aqui a análise feita é do quadro geral de indicadores e sua correlação.

A correlação de Pearson indica associação entre variáveis, ou seja, variação compartilhada dos escores dessas variáveis (Figueiredo Filho e Silva Jr., 2009). Quando dois ou mais indicadores apresentam previsibilidade entre si, ou seja, sabendo a resposta de um podemos antecipar a resposta do outro, há uma relação entre eles.

A matriz de correlação entre as variáveis/indicadores pode ser executada através dos códigos:

### **Código no SPSS:**

Menu Analyze > Correlate > Bivariate > Variables:v1+v2+v3

### **Código no Stata:**

```
correlate v1 v2 v3
```

### **Código no R:**

```
cormat <- cor(dados)
round(cormat, 2)
```

Uma correlação entre os indicadores pode denotar que eles pertencem a uma mesma dimensão, tornando mais fácil agregar os dados em fatores ou componentes (ver tópico 3 a seguir). Análise de correlações também podem indicar quais variáveis serão redundantes no modelo (Sullivan, 1974).

Pode ser tentador para o pesquisador incluir diversos indicadores no índice, e assim tornar difícil para si próprio a tarefa de eliminar alguns deles. Por isso é bom sempre ter em mente que esses indicadores, na pura semântica da palavra, apenas indicam o conceito em formato quantificado, não o substituindo (Jannuzzi, 2002). Sabendo que o conceito terá uma representação matemática no indicador, não devemos supervalorizá-lo na construção do índice.

## **4.3 Pontuação Do Índice**

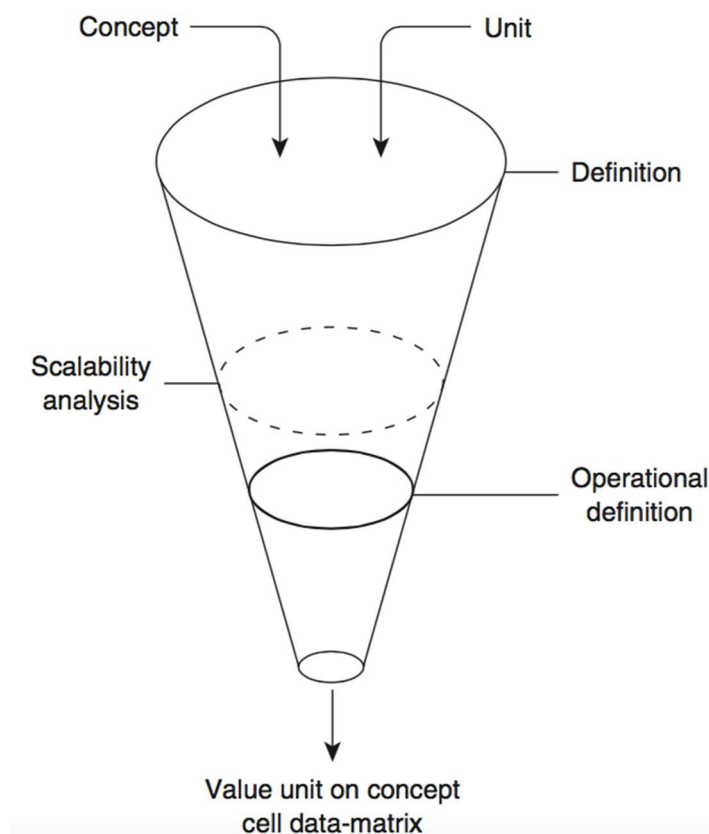
O leitor logo irá descobrir que não há meio-termo para esta etapa, ou é uma tarefa divertida para o pesquisador, ou completamente insuportável. Por que isso? Ora, atribuir números concretos a conceitos políticos e sociológicos tem cunho bastante arbitrário. Mesmo dor, exemplo usado no começo do artigo, não é tão palpável valorativamente. A operacionalização consiste em transformar conceitos teóricos em variáveis observadas, seja direta, seja indiretamente.

É importante saber diferenciar o conceito de sua mensuração. Um conceito pode ser contínuo, ou seja, varia em nível, e o pesquisador pode mesmo assim adotar uma mensuração dicotômica para ele (Bollen, 1990). Para o caso de democracia, uma mensuração dicotômica teria apenas as opções de valores “democrático” e “não-democrático”. Um índice ajuda a mensurar o conceito de forma contínua.

Ao pontuar o índice (*index scoring*) estamos operacionalizando o conceito. É o momento ápice da mensuração. Pennings, Keman e Kleinnijenhuis (2006) denominaram esse procedimento de “funil da operacionalização”. Na boca larga do funil, que é a definição, colocamos o conceito e

a unidade. Eles vão escorregando pelo primeiro canal do estreitamento, que é a análise de escalabilidade, depois pelo segundo canal do estreitamento, que é a definição operacional, até chegar no bico mínimo do funil, cujo resultado é o valor unitário numa matriz de células conceituais.

**Figura 3 - "funil da operacionalização"**



Fonte: Pennings, Keman e Kleinnijenhuis (2006)

Há dois potenciais problemas que surgem na atribuição de valores. Um deles vem se manifestar quando vamos aplicar esses valores aos casos. O segundo problema é em relação ao peso dado aos indicadores. Veja abaixo:

**Problema 1:** Como classificá-los? Por que determinado caso tem valor 1 ou 2 quando poderia ter 1,5? Quanto a isso não há regra fixa e cabe ao pesquisador justificar sua escolha, consciente de que inevitavelmente será criticado.

**Solução:** A análise fatorial é uma técnica estatística de redução de dados e identificação da estrutura dos mesmos através das inter-correlações das variáveis (Hair et al, 2005). Esse processo é feita pelo cálculo de escores que substituem os valores originais das variáveis (Hair et al, 2005),

assim como a atribuição citada acima, porém a análise fatorial o faz com base na estrutura real dos dados, e não com arbitrariedade.

Para exemplos aplicados de análise fatorial, e como selecionar o número de fatores, há os trabalhos de Bezerra e Corrar (2006) e Figueiredo Filho et al (2014). Os códigos básicos iniciais para utilização desta técnica estão a seguir:

**Código no SPSS:**

```
Menu Analyze > Dimension Reduction > Factor
```

**Código no Stata:**

```
factor v1 v2 v3 ou factor v1-v3
```

**Código no R:**

```
fit <- factanal(dados, factors=x, rotation="varimax")  
print(fit)
```

**Problema 2:** Para ilustrar, o pesquisador pode decidir que um indicador classifique seus casos entre 0 e 1, enquanto que outro indicador o faça entre 0 e 2. Isso significa dar pesos diferentes aos indicadores.

**Solução teórica:** Babbie (2010) argumenta que apesar dessa decisão caber ao pesquisador, a melhor estratégia é pesar igualmente os indicadores.

**Solução empírica:** Há algumas estratégias estatísticas para avaliar a forma de pesar variáveis, e mesmo excluir ou agregar algumas do índice. A Análise de Componentes Principais, uma progressão da análise fatorial, permite agrupamento de informação em fatores com perda mínima de dados originais, e também a possibilita identificar quais desses fatores contribuem mais para explicar de sua variância (Hair et al, 2005).

Figueiredo Filho et al (2013) traz o uso de análise de componentes principais na construção de indicadores sociais. Mendes (2014) aplica a técnica para seu índice de desenvolvimento institucional de municípios. Por último, uma aplicação clássica é a de Putnam (2000) em seu índice de capital social.

Para realizar uma análise de componentes principais básica você pode usar os códigos:

**Código no SPSS:**

```
Menu Analyze > Dimension Reduction > Factor > Extraction >  
Method:Principal Components
```

**Código no Stata:**

```
pca v1 v2 v3
```

**Código no R:**

```
dados.pca <- prcomp(dados, center = TRUE, scale. = TRUE)
```

#### 4.4 Validação Do Índice

A última etapa é a validação do índice - mostrar que ele de fato mede o que se propõe a medir. Já discuti aqui o que significa validade e seu papel na construção de um índice. Isso é diferente do processo de validação da mensuração, e não se deve confundir os termos (ADCOCK & COLLIER, 2001). Basicamente há duas verificações a serem feitas, em nível micro e nível macro.

Em nível micro, temos que verificar se um indicador está contribuindo para o índice individualmente ou apenas reproduzindo as informações de outro indicador presente (Babbie, 2010). Inversamente, um indicador pode simplesmente não refletir por completo o conceito desejado. É preciso buscar um equilíbrio entre a parcimônia proposta pela primeira situação e a inclusão de componentes que pode surgir da segunda situação (ADCOCK & COLLIER, 2001).

Em nível macro, sabemos que quando um estudo pode ser generalizado para além dos casos observados, ele tem validade externa (DE VAUS, 2001). O processo de validação externa de um índice acontece quando outros índices - logo, diferentes mensurações - chegam às mesmas conclusões que o seu (BABBIE, 2010). Por exemplo, existem diversos índices para democracia, o que é excelente para a ciência, mas se dois índices concluem divergentemente se um país é democrático ou não, isso se torna problemático.

Outro caso é o de avaliação de aprendizado de estudantes. Em 2015, o IDEB colocou Pernambuco em primeiro lugar no desempenho de alunos do ensino médio público. Já no PISA 2015 (Programa Internacional de Avaliação de Estudantes), Pernambuco figura em décimo-nono dentre os estados brasileiros. Onde está o problema? Na seleção de indicadores, em sua mensuração, na pontuação de itens? Devido à divergência, há possibilidade de que um dos índices não seja válido.

Visando trabalhar com esta questão, pode ser feita uma análise quali-quantitativa entre índices existentes. A priori uma análise qualitativa vai a fundo na construção de cada índice (selecionado

para análise - é importante que eles tenham os mesmos casos) para verificar se eles mensuram as mesmas coisas.

A análise quantitativa envolve alguns testes estatísticos, como a já citada Análise de Componentes Principais, e também correlação entre pares de índices, como executado por Lorenz (2010) ao comparar diferentes índices de rigidez constitucional. Uma técnica que mede a confiabilidade do índice é o alfa de Cronbach, através da consistência da escala inteira, sendo o limite para pesquisas exploratórias de 0,60 (Hair et al, 2005).

Para calcular o alfa de Cronbach:

#### **Código no SPSS:**

Menu Analyze > Scale > Reliability Analysis > Items:v1 v2 v3 > Model:Alpha

#### **Código no Stata:**

```
.alpha v1-v3, item
```

#### **Código no R:**

```
library(psych)  
psych::alpha(dados)
```

A convergência de associações empíricas entre indicadores em geral é esperada, e quando há divergência o pesquisador pode retornar ao conceito para verificar se: sua sistematização foi feita de forma incorreta, ou se apenas o conceito sistematizado é diferente daquele indicador que se buscou associação (Adcock e Collier, 2001). Um outlier também pode estar perturbando as análises: explore bem os seus dados!

### **4.5 Demais Questões De Análise**

Por fim, alguns pontos a serem considerados durante a construção de índices. Um deles é um problema recorrente, o de dados inexistentes. Na literatura específica o leitor encontrará este termo com a denominação de *missing data*. É comum que haja observações em branco na matriz, seja de forma sistemática para determinadas variáveis ou casos, quando não existem os dados, ou aleatoriamente, quando o dado se perdeu o deixou de ser preenchido (em caso de questionário por exemplo).

Qualitativamente algumas escolhas podem ser feitas. Para poucos casos com dados inexistentes, eles podem ser excluídos da análise. Também pode-se optar por preencher com uma das respostas disponíveis com base nas outras - isso é comum em questionários - ou com o meio-termo, em caso de pontuação. Por último, também é possível simplesmente interpretar o que a inexistência de dado pode significar e incluir isso na análise (Babbie, 2010).

Quantitativamente, mais uma vez recorrendo à estatística, uma saída é verificar se o indicador que apresenta dados faltando tem correlação com outro indicador, e utilizar-se disso para prever os valores incompletos e preenchê-los (Little e Rubin, 1987). É preciso tomar cuidado para não cair na ilusão de que esses valores não foram criados artificialmente. Uma forma mais refinada dessa tática é o método Buck, que diante de distribuição normal multivariada numa matriz de covariância, estima os coeficientes de regressão das variáveis incompletas a partir da função gerada pelos casos completos. Tudo isso gera algum nível de distorção, e não há método que funcione com perfeição nesse tipo de caso (Little e Rubin, 1987).

Tentativas mais recentes de aprimorar esse processo chegam na técnica de *multiple imputation*, com base em Little e Rubin (1987) e trabalhos anteriores. Os valores são imputados nas células vazias através de um modelo paramétrico que assume que os dados completos possuem distribuição probabilística conjunta (Lall, 2016).

O uso desta técnica leva em conta os problemas que surgem da simples exclusão de dados inexistentes, especialmente na Ciência Política, onde essa ausência pode ser não-aleatória, causada por desenvolvimento econômico ou instituições políticas (Lall, 2016).

Um segundo ponto vai além da construção do índice para a seleção de casos que receberão o tratamento do índice. Lorenz (2010) se questiona: como escolher os países? Afinal, ao comparar índices temos de usar o mesmo conjunto de casos. Isso levanta a preocupação com o viés de seleção. Ao partir para a análise de um índice com viés, um indicador pode mostrar ter um efeito causal que não existe. Deve-se evitar os erros comuns que levam a viés: selecionar pela variável dependente, selecionar de acordo com um ponto fixo no tempo, e selecionar com base em *path-dependence* (Geddes, 1990).

Por último, vem-se desenvolvendo técnicas cada vez mais sofisticadas para a construção de índices. Tomando especificamente o exemplo de índices de desenvolvimento social, Foe e Tanner (2012) propõem utilizar *Matching Percentiles* para agregar variáveis, com as vantagens de que este método lida melhor com os valores inexistentes, com o processamento de variáveis binárias (sim e não, certo e errado, presente e ausente), com poucos países no conjunto de casos, e também dispensa assumir uma distribuição linear dos indicadores.

Uma recente empreitada também é bastante interessante: Souza, Graça e Silva (2017) usam análise de rede de seguidores de deputados federais no Twitter e inferência bayesiana para estimar pontos ideais (escores) na escala de ideologia política.

Este tópico serve como um lembrete ao leitor de que apesar dos critérios elementares, devemos ficar atentos aos diversos detalhes a serem examinados. Além disso, os métodos científicos estão sempre evoluindo em busca de resultados melhores e mais confiáveis.



## 5. CONSIDERAÇÕES FINAIS

O leitor abriu este artigo preocupado com o uso de índices nas ciências sociais. É uma preocupação pertinente. Procurei aqui mostrar que, na ciência, as inquietações devem ser resolvidas com critérios rigorosos, deixando visíveis as falhas presentes e trabalhando para amenizar seus efeitos. Isso não se mostra diferente nas ciências sociais, apesar da complexidade de seus objetos.

Os índices muitas vezes são inevitáveis, pela multidimensionalidade que comentei no princípio. Portanto, aprimorar sua construção, como sugeri na última seção, é essencial. Não há um único método correto, e devemos nos aproveitar das divergências entre pesquisadores para alavancar as discussões nas ciências sociais. Este artigo buscou trazer diferentes perspectivas e sugestões sobre o caminho a ser seguido.

É importante destacar que dificuldades encontradas por pesquisadores e mesmo estudantes não são aleatórias. Por exemplo, no trabalho de Souza, Graça e Silva (2017) os autores comentam sobre o uso de computadores de alto processamento de dados em trabalhos semelhantes que eles tentam replicar. Ressaltam que nas ciências sociais, e principalmente no Brasil, esse tipo de computador não é de comum acesso, resultando em maior emprego de tempo para executar as mesmas tarefas.

Questões desse tipo não são triviais, e nem de fácil solução. A maior parte sequer está ao alcance do pesquisador. Apesar da limitação de recursos, investir no aprendizado metodológico é uma importante munição para o desenvolvimento da ciência. O acesso a dados e tecnologias pode ser escasso comparativamente, mas os trabalhos produzidos serão rigorosos e válidos na construção do conhecimento científico. Artigos metodológicos de toda natureza são bem-vindos na Ciência Política para se inserir nesse objetivo.

Como venho insistindo desde o começo deste *paper*, o uso de técnicas de pesquisa deve sempre vir acompanhado de boas noções de teoria de mensuração. Não é à toa que o primeiro passo para construção de índices está na delimitação empírica do conceito. Conceitos complexos são ótimos para debate teórico, mas sem mensuração adequada, perde-se o desenvolvimento empírico da disciplina. Essa recomendação fica não só para os estudantes que estudam como criar índices, mas também aos professores para que incluam teoria de mensuração em seus cursos de metodologia.

Quais temas e conceitos você, leitor, acha que ainda são sub-explorados na Ciência Política ou outra área de conhecimento? É possível que esta introdução ao uso de índices tenha lhe ajudado, ou mesmo incentivado, a suprir essa falta? Ou você acha que alguma mensuração existente não está

sendo feita da forma correta, e se propôs a replicar usando métodos mais rigorosos? Se as respostas forem positiva, o artigo cumpriu o seu papel.

## REFERÊNCIAS BIBLIOGRÁFICAS

ADCOCK, Robert and COLLIER, David. (2001) "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95:3 (September), 529-46.

BABBIE, Earl. (2010) "Indexes, Scales and Typologies." In: *The Practice of Social Research*. Twelfth Edition. Wadsworth, Cengage Learning.

BEZERRA, Francisco Antonio, and Luiz J. CORRAR. (2006). "Utilização da análise fatorial na identificação dos principais indicadores para avaliação do desempenho financeiro: uma aplicação nas empresas de seguros." *Revista Contabilidade & Finanças* 17, no. 42: 50-62.

BOLLEN, Kenneth A. (1990). "Political democracy: Conceptual and measurement traps." *Studies in Comparative International Development* 25, no. 1: 7-24.

BRADY, Henry and COLLIER, David. (2004) *Rethinking Social Inquiry*. Rowman and Littlefield.

CUKIERMAN, Alex; WEBB, Steven B.; NEYAPTI, Bilin. (1992) "Measuring the Independence of Central Banks and Its Effect on Policy Outcomes." *The World Bank Economic Review*. Vol. 6, No. 3: 353-398.

ESI Triage Research Team (2004). Disponível em: <<http://www.esitriage.org/index.asp>>

DAHL, Robert. (1997) *Poliarquia: participação e oposição*. 1ed. 3 reimp. Coleção Clássicos, EDUSP. São Paulo.

DEMOCRACY BAROMETER (2016). Disponível em: <[www.democracybarometer.org](http://www.democracybarometer.org)>

DE VAUS, David (2001). "Tools for research design." In *Research Design in Social Research*, ed. David De Vaus. Thousand Oaks, CA: Sage.

FIGUEIREDO FILHO, D. B. ; SILVA JÚNIOR, J. A. (2009). "Desvendando os Mistérios do Coeficiente de Pearson (r)." *Política Hoje* (UFPE. Impresso), v. 18, p. 115-146.

FIGUEIREDO FILHO, Dalson Britto, Enivaldo Carvalho da ROCHA, Ranulfo PARANHOS, Anderson Henrique SILVA, José Alexandre SILVA JR, Lucas SILVA, and Dáfni Priscila ALVES. (2014). "Análise Fatorial Garantida ou o Seu Dinheiro de Volta: Uma Introdução à Redução de Dados." *Revista Eletrônica de Ciência Política* 5, no. 2.

FIGUEIREDO FILHO, Dalson Britto, Ranulfo PARANHOS, Enivaldo Carvalho da ROCHA, and Romero G. MAIA. (2013). "Análise de componentes principais para construção de indicadores sociais." *Rev. Bras. Biom* 31, no. 1: 61-78.

FOE, R, & TANNER, J.C. (No. 2012-04) "Methodology of the Indices of Social Development". *ISD Working Paper Series*. Retrieved from <<http://hdl.handle.net/1765/50510>>

FREEDOM HOUSE (2016). Disponível em: <<https://freedomhouse.org>>

GEDDES, Barbara. (1990) "How the cases you choose affect the answers you get: selection bias in comparative politics". *Political Analysis*. 2 (1): 131-150.

GILL, J. (2006). *Essential mathematics for political and social research* (pp. 271-277). Cambridge: Cambridge University Press.

HAIR Jr., Joseph F.; BLACK, William C.; BABIN, Barry J.; ANDERSON, Rolph E. (2005) *Análise Multivariada de Dados*. Quinta edição. Bookman, Porto Alegre.

HANSEN, Jaqueline Resmini. (2016) "*No computador, na rua ou no smartphone: condicionantes ou covariantes do engajamento online*." PhD diss., Dissertação de mestrado, Programa de Pós-Graduação em Ciências Sociais, UEM, Maringá.

JANNUZZI, Paulo de Martino. (2002) "Considerações sobre o uso, mau uso e abuso dos indicadores sociais na formulação e avaliação de políticas públicas municipais." *RAP*, 36(1):51-72. Rio de Janeiro.

KELLSTEDT, Paul M; WHITTEN, Guy D. (2009) *The Fundamentals of Political Science Research*. Cambridge University Press.

KING, Gary; KEOHANE, Robert and VERBA, Sidney. (1994) *Designing Social Inquiry: scientific inference in qualitative research*. Princeton University Press.

KUBRUSLY, Lucia Silva. (2001) "Um procedimento para calcular índices a partir de uma base de dados multivariados." *Pesquisa Operacional* 21, no. 1: 107-117.

LALL, Ranjit. (2016). "How multiple imputation makes a difference." *Political Analysis* 24, no. 4: 414-433.

LANDMAN, Todd. (2008) *Issues and Methods in Comparative Politics: an introduction*. Third Edition. Routledge, Taylor & Francis Group e-Library.

LIJPHART, Arend. (2008) *Modelos de democracia: desempenho e padrões de governo em 36 países*. 2a. ed. Rio de Janeiro: Civilização Brasileira.

LITTLE, Roderick and RUBIN, Donald B. (1987) *Statistical Analysis with Missing Data*. New York: John Wiley & Sons.

LORENZ, Astrid. (2005) "How to Measure Constitutional Rigidity: Four Concepts and Two Alternatives." *Journal of Theoretical Politics*, 17(3): 339–361. SAGE.

MAHONEY, James e VANDERPOEL, Rachel Sweet. (2014) "Set Diagrams and Qualitative Research." *Comparative Political Studies*, 2015 48: 65. SAGE Publications.

MAINWARING, Scott, Daniel BRINKS, and Aníbal PÉREZ-LIÑÁN. (2001) "Classifying political regimes in Latin America, 1945-1999." *Dados* 44, no. 4: 645-687.

MENDES, Renata. (2014) "Brazilian National Housing Policy: Institutional capacity and housing deficit." *Politikon: IAPSS Political Science Journal*, Vol. 23, June.

MERKEL, Wolfgang and BOCHSLER, Daniel (project leaders) (2016); Bousbah, Karima; Bühlmann, Marc; Giebler, Heiko; Hänni, Miriam; Heyne, Lea; Müller, Lisa; Ruth, Saskia; Wessels, Bernhard. *Democracy Barometer. Methodology. Version 5*. Aarau: Zentrum für Demokratie.

MOORE, W. H.,; SIEGEL, D. A. (2013). "*A mathematics course for political and social research*". Princeton University Press.

MOZZAFFAR, S., & SCHEDLER, A. (2002) "The Comparative Study of Electoral Governance—Introduction." *International Political Science Review*, 23, 5–27.

NUNNALLY, Jum C., and Ira H. BERNSTEIN. (1978) "*Psychological theory*." New York, NY: MacGraw-Hill.

PENNINGS, Paul; KEMAN, Hans; KLEINNIJENHUIS, Jan. (2006) *Doing Research in Political Science*. 2nd edition. SAGE Publications. London.

POGREBINSCHI, Thamy. (2010) "Democracia Pragmática: Pressupostos de uma Teoria Normativa Empiricamente Orientada." *DADOS – Revista de Ciências Sociais*, Rio de Janeiro, vol. 53, n 3, pp. 657 a 693.

PUTNAM, Robert D. (2000). *Bowling alone: America's declining social capital*. Simon & Schuster, New York.

ROBERTS, Fred S. (2009) *Measurement Theory*. Cambridge University Press (1985), digitally printed version.

SANTOS, B. R. L., TEODORO, J. V., & FACCIN, A. C. T. M. (2015). "Análise de componentes principais dos indicadores de qualidade de vida". *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, 3(1).

SARTORI, Giovanni. (1970) "Concept Misinformation in Comparative Politics." *American Political Science Review*. 64:4 (December), 1033-53.

SOUZA, Rafael Martins de, Luís Felipe Guedes da GRAÇA, and Ralph dos Santos SILVA. (2017) "Politics on the Web: Using Twitter to Estimate the Ideological Positions of Brazilian Representatives." *Brazilian Political Science Review* 11, no. 3.

STRAUSS, Leo. (2009) *Direito Natural e História*. - Reimp. (Biblioteca de teoria política; 3). Edições 70. Lisboa.

SULLIVAN, John L. (1974). "Multiple indicators: Some criteria of selection." In *Measurement in the social sciences* / Hubert Blalock (ed.), pp. 243-269. The Macmillan Press.

TAROUÇO, Gabriela. (2012) "Governança eleitoral: incentivos institucionais e competição partidária". *8o Encontro da Associação Brasileira de Ciência Política*. Gramado.

VAN EVERA, Stephen. (1997) "Hypotheses, laws and theories: a user's guide." In Stephen Van Evera. *Guide to methods for students of Political Science*. Cornell University Press.

WEBER, Max. (1982) "A Ciência como Vocação". In: *Ensaio de Sociologia*. LTC – Livros Técnicos e Científicos Editora S.A.

# **On A Scale Of 1 To 10, How Much It Hurts? On The Construction Of Indexes And Their Validity In Social Sciences**

Haína Coelho. Political Science Ph.D Candidate at UFPE. E-mail: haina\_sama@hotmail.com

## **Abstract**

The goal of this paper is to provide a methodological guide for elaboration of indexes, or composite indicators. I clarify about types of indexes according to their data measurement, and I discuss about its use in the social sciences employing the debate about normativity on science. I give introductory notions on measurement theory, and its use on indexes elaboration. Each step to build a valid index is detailed: from item selection, going through empirical relation analysis, item scoring and finally validation of the index, eventually followed by execution codes for the techniques in softwares SPSS, Stata and R. To conclude, I raise some extra questions that could appear on the way to build an index in the social sciences, such as missing data.

**Keywords:** Indexes; Validity; Measurement; Social Sciences.

## **Resumo**

Este artigo fornece um guia metodológico para a elaboração de índices. Esclareço sobre os tipos de índices de acordo com o nível de mensuração das variáveis e faço uma discussão sobre a aplicação de indicadores compostos nas ciências sociais utilizando o debate de normatividade na ciência. São dadas noções iniciais sobre teoria de mensuração, e sua utilização na construção de índices. Cada passo para criar um índice é detalhado: desde seleção de itens, passando por análise de relações empíricas, ponderação dos itens e validação do indicador, acompanhados eventualmente de códigos de execução das técnicas nos softwares SPSS, Stata e R. Para finalizar, são levantadas alguns obstáculos que podem aparecer no caminho para a construção de índices nas ciências sociais, como dados inexistentes.

**Palavras-Chave:** Índices; Validade; Mensuração; Ciências Sociais.

# On A Scale Of 1 To 10, How Much It Hurts? On The Construction Of Indexes And Their Validity In Social Sciences

Haína Coelho – Universidade Federal de Pernambuco

*"A major difference between a "well-developed" science such as physics and some of the less "well-developed" sciences such as psychology or sociology is the degree to which things are measured."*

*(Fred Roberts)*

## 1. INTRODUÇÃO

In hospital emergencies, in order to highlight the priority of care, the patient undergoes a screening where their risk classification is made. This is based on some indicators. For example, there is an index that is internationally used for hospital risk assessment, the Emergency Severity Index. It considers whether the patient needs immediate intervention, whether they are in pain or at risk of worsening, and whether additional resources for consultation - such as laboratory tests - will be needed (ESI Triage Research Team, 2004). There are other indexes used by medicine with this same purpose.

As the attentive reader may deduce, although the data is primarily produced by screening, it is purely observational. After all, it is impossible to estimate exactly how much pain someone else is feeling, no matter how much they describe it. This can lead to errors of measurement by the nursing, classifying a patient in a wrong way in the scale of risk. The indexes in the social sciences work in a similar way.

Indexes are composite ordinal measures (BABBIE, 2010), that is, they measure a concept in gradation of levels from the scores of two or more variables that compose this concept. Scores can be either continuous or dichotomous. The scores of all variables are aggregated into a single measure, the index.

Some of these indexes are very famous. The Human Development Index (HDI) is one of them, and it calculates human development levels of countries through four indicators: literacy rate, schooling rate, life expectancy, and GDP per capita. In a similar way, there are studies seeking to aggregate for analysis highly disseminated indicators of quality of life, such as life expectancy and nutrition of the population (SANTOS, TEODORO & FACCIN, 2015).

Indexes that measure levels of education, such as the IDEB (Basic Education Development Index), are useful as a reference to public officials and managers in improving public policies.

A possible obstacle to the creation of indexes in the social sciences lies in the complexity and fluidity of concepts, and especially in the difficulty of measuring them. Theory of measurement is a necessity to the empirical development of any discipline, and where is one of the great calluses of the social sciences.

This paper aims to provide an introductory guide to the construction of indexes and basics of measurement theory. The target audience is comprised of undergraduate and graduate students from all areas, but focusing more on Social Sciences students, especially Political Science. A beginner in scientific methodology may use this article as a basic reference for index building, and they will be directed to sources that deepen their specific needs.

Systematically disseminating the knowledge needed to improve the transformation of theoretical concepts into tangible and testable measures is the best way to promote the empirical development of our area of knowledge. The article, trying to reach the individuals in the initial stage, is a mission in this sense.

In the following section, the nature of an index and its uses will be clarified. The subsequent section will briefly discuss the validity of this technique in the social sciences, and finally, I bring the steps for the proper construction of an index, before the final considerations.

## **2. INDEXES: WHAT ARE THEY, WHAT ARE THEY FOR AND HOW ARE THEY CALCULATED?**

To understand what are indexes I follow a logical step-by-step, starting at the beginning, so I ask the bored reader a little patience. The theory is almost always the starting point in scientific research, and it is from its development that we extract a series of abstract concepts (VAN EVERA, 1997). When values are assigned to these concepts, they become the research variables (VAN EVERA, 1997).

However, the interesting part of science is to compare variables and to establish connections between them, especially those of a causal nature. For this to be possible, we enter into mathematical relations (MOORE & SIEGEL, 2013).

A relation between concepts exists when from one variable we can obtain information about another variable, and thus better describe the universe we are observing (Moore and Siegel, 2013). It is important, then, that we know how to accurately represent these relations, or we would be performing a bad explanation of reality.



Measurement is the crucial step for this, since we assign the most adequate values to preserve the relations between variables - which in math becomes a manipulative function (Roberts, 2009). The measurement represents quantities, or it defines in which categories the attributes of certain objects are framed (NUNNALLY & BERNSTEIN, 1978). The two variations respectively linked to the measurement of directly measurable attributes, such as height, and abstract categories, such as intelligence (NUNNALLY & BERNSTEIN, 1978).

Any scientist develops measures that obey certain rules for quantifying attributes and uses mathematical systems to analyze the structure of the data. (NUNNALLY & BERNSTEIN, 1978). This differs from purely mathematical measurement by the connection with the empirical universe: a mathematical system only needs internal validation of rules to make sense (NUNNALLY & BERNSTEIN, 1978).

The process of measurement, mathematical treatment of concepts, is what allows Political Science to transform its concepts into variables as such, and not only to categorical classifications (SARTORI, 1970). The attributes belonging to the concepts must be capable of gradation, which is different from naively assigning numerical values to categories (SARTORI, 1970).

There is no single way to measure concepts, as this depends on the type of data the research is dealing with. Simplistically, if the data has categorical values, they are called discrete. But if values are real numbers, they are continuous data. For discrete data, the measurement can be nominal, when it is purely categorical, or ordinal, where there is hierarchy between categories. For these two types of data, we have the interval measurement, where the space between values is always the same, and ratio measurement, executed in the same way as the interval, but considering the zero point as an indicator of absence of effect (Gill, 2006). Table 1 summarizes this information.

**Table 1 – Types of indexes**

Measurement level	Data type	Example
<b>Nominal</b>	discrete	War/peace
<b>Ordinal</b>	discrete	Positioning on the ideological scale
<b>Interval</b>	discrete or continuous	Voter perception of corruption
<b>Ratio</b>	discrete or continuous	Number of vetoes by the Executive

**Source:** Author's elaboration, compiled based on Gill (2006), Moore and Siegel (2013), Roberts (2009).

In the social sciences, highlighting Political Science, the measurement process is more complex because of the multidimensional nature of many concepts. For example, the concept of political participation, whose attributes open the possibility of various forms of participation (HANSEN, 2016).

This aspect leads the researchers to identify indicators for these concepts, adopting the most diverse strategies to measure them, be it dichotomous, polycotomic, scalar intervals, etc. (LANDMAN, 2008).

The indexes, then, arise from the multivariate measurement of these indicators (HAIR *et al*, 2013). It is important to differentiate indexes and scales, although this factor is not so relevant to this article. A scale uses the intensity differences in the indicators of a variable to identify patterns, while an index accumulates the values of each indicator to rank the observations more generally (BABBIE, 2010).

The use of indexes and scales is recurrent in Political Science. Democracy studies are quite self-evident, for example, the work of classifying regimes (MAINWARING, BRINKS & PEREZ-LIÑAN, 2001; DAHL, 1997), democracy quality (Democracy Barometer, 2016; Freedom House, 2016), and we can also mention studies on institutional designs, such as central bank independence (LIJPHART, 2008; CUKIERMAN, WEBB & NEYAPTI, 1992), constitutional rigidity (Lorenz, 2005; Lijphart, 2008), and electoral governance (MOZZAFFAR & SCHEDLER, 2002; TAROUCO, 2012).

The first point highlighted in the article was the difficulty nurses have in measuring a data that cannot be directly observed, and that this is repeated in the social sciences. Gill (2006) warns how often it is the construction of artificial indexes by these researchers to measure some characteristics of society through ordinal data. What is the real nature of this problem? The next section looks at this question.

### **3. CAN WE USE INDEXES?**

As stated in the previous section, mathematics turns the relation among variables into functions so that they can be manipulated and compared. However, some data can't be manipulated, such as the symptoms of a patient in the emergency, the institutional design of a country, the characteristics of a democratic regime, or the implementation of a public policy. This is the reality for many information considered relevant in the social sciences.

This is one of the great methodological questions presented by Brady and Collier (2010): the dichotomy between observational data and experimental data. When working with observational

data, not having the ability to manipulate them means not being able to control its effect on the analyzed relationship. Observational data are often also not objective to measure.

This factor raises one of the major trade-offs of scientific research: reliability versus validity. Reliability exists when different measurements of the same phenomenon generate the same results, and validity when the measurements reflect what the researcher wants to measure (King, Keohane and Verba, 1994). From now on I use democracy as an example, due to its iconic applicability. However, the concept could be X, Y or Z, so if some other is more logical for you, do the mental exercise of replacing the terms.

Thinking about democracy, if we choose a polychotomous concept and want to compare between countries, or a time series from a single country, we have to choose indicators that accuse a certain democratic level in an observation every time we measure this observation, and these indicators should in fact indicate the democratic level.

It is the controversy referred to by Kellstedt and Whitten (2009) when they state that no matter how tempting it is to think of democracy as a pregnancy (it is or is not), it functions as a continuum of levels, and the difficult part is to identify them. In addition to the natural existence of the tradeoff, the multidimensional concept of democracy, and the non-manipulability of its indicators, make this task quite complex.

At this point the reader wonders whether such an effort is worth it, or whether constructed indexes will mean anything scientifically. To answer the first reflection, we will return to the primary debate on normativity in Political Science and ask the reader to follow another digression.

Weber (1982), in a sense, asked the same question. He did not refer to specific indexes, obviously, but to the understanding and interpretation of political phenomena in general. Science does not respond how these objects are compensating enough for the effort to know them. The value of a political position embedded in it must never be carried into the room, for it runs the risk of turning science into demagoguery (WEBER, 1982 [1946]).

This argument in no way helps the reader's first reflection, perhaps reinforcing it. But in dialogue these Weberian ideas, Leo Strauss (2009 [1953]) brings a new interpretation. He points out how the position of differentiation of facts and values is not a rejection of the is/ought but to reject that it is possible to know the is/ought, and to explain a social phenomenon we must know it for what it is. What Strauss claims, however, is that it is difficult to understand a social phenomenon without a prior conceptual parameter. When we reject value judgments, we are avoiding "calling things by their names" as he puts them, and we even threaten the objectivity of interpretation. The objects of the social sciences, to close their debate with Weber, arise through the reference of values. If we return to our example of the quality of democracy, this becomes more evident.

What Strauss is suggesting is what later motivates the analytical and epistemological debate of pragmatic democracy: the dichotomization between normativity and empiricism by contemporary Political Science must be overcome. It is possible to endorse values without giving up facts, and it is useful for discipline, because of its nature, to "reconcile the empirical need to explain political institutions with the normative ambition to propose more desirable forms of political organization" (POGREBINSCHI, 2010, p. 689).

That is, Weber's compensation for the study of social phenomena lies in normativity: not in the is/ought of the object, but in its choice - that is why we still study the quality of democracy, despite the prickly path and severe criticism shall be taken in that respect. The use of indexes accompanies this logical result due to the aforementioned multidimensional character of the concepts that make it necessary.

Although we can't measure the democratic efficiency of an institutional design such as the performance of a car, it is important to differentiate between what works and what does not. This brings us to the second reflection of the reader: do these analyzes have any use? To this I answer: they may have, if they follow a series of strict criteria. The next section develops these criteria in a way that attempts to overcome problems with indexes in the social sciences.

#### **4. How to create a valid index?**

Although several researchers have worked on this theme, even if in a scattered way, I will here systematize the process of index construction, reproducing the four steps described by Babbie (2010) - item selection, empirical relationship analysis, index score, and its validation - and deepening the methodological discussion in each of them.

##### **4.1 Item Selection**

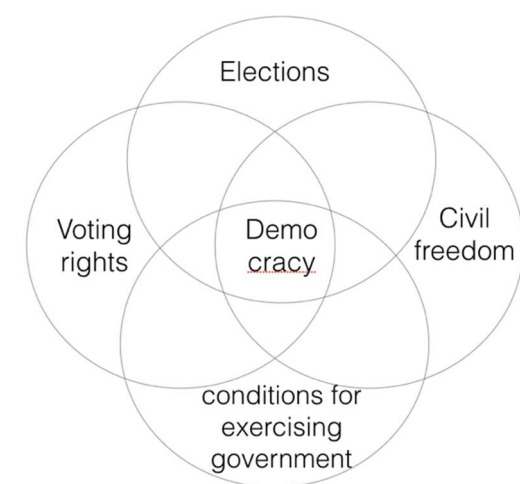
It is important to clarify that item selection refers to the components of the index, not the research cases. The first step is directly linked to the use of theoretically oriented concepts. For the index to represent the concept we want, we must choose variables that preserve the meaning of the concept. For this, the concept must be well delimited. Very abstract concepts are of little use in indexes since they make difficult the operationalization of the variables.

For example, how to measure the concept of alienation? By the way, what is alienation? One of the problems of concepts in Political Science is that they often denote broadness enough to undermine its empirical precision, which is the so-called conceptual stretching (SARTORI, 1970).

Conceptual stretching is different from the generalization of a concept by the diminution of attributes, so that it is more inclusive: it is about leaving the concept broad but without many specificities. That is, vague and obscure (SARTORI, 1970).

It is necessary that the concept be well defined and whether the indicators refer to it in general or to its dimensions in a specific way (Babbie, 2010). One way to work these crucial decisions clearly is to use Venn diagrams. For the concept in general, select the necessary and sufficient conditions that form it (MAHONEY & VANDERPOEL, 2014). These conditions can already be interpreted as the indicators, if operationalized in this way, and visually the diagrams characterize a relation of intersection of sets, as in Figure 1 below:

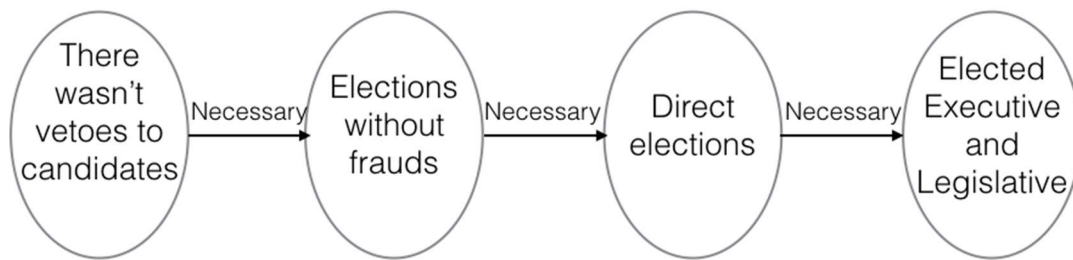
**Figure 1 - Necessary And Sufficient Conditions For Democracy**



Source: author's elaboration, based on Mainwaring, Brinks, Perez-Liñan (2001).

For the case of a specific dimension, the method of elaboration in sequence works well. It organizes the sets into a chain of necessary or sufficient conditions (Mahoney and Vanderpoel, 2014). The logic is that for each dimension there may be a gradation. Continuing with the example of democracy, I can choose a dimension in Figure 2:

**Figure 2 – Necessary conditions for the Elections dimension**



Source: author's elaboration, based on Mainwaring, Brinks, Perez-Liñan (2001).

By ensuring that the indicators represent a single dimension each, we avoid that more general concepts have repeated dimensions between them. This tactic can also act as a way to separate the endogenous components from exogenous, for which King, Keohane and Verba (1994) warn. It is also good to guarantee variance within the indicators, because if one of them is constructed in such a way that all cases or no case is included, it will not be useful for the index (BABBIE, 2010). An empirical example of high refinement of item selection for index construction is the Democracy Barometer's conceptual tree (MERKEL, BOCHSLER, ET AL, 2016).

A purely statistical item selection strategy is clusters analysis. This technique identifies which variables do not change the results if they are removed from the analysis by measures of similarity between objects (KUBRUSLY, 2001).

For a basic clusters analysis, the following software codes can be used:

**SPSS code:**

```
Menu Analyze > Classify > K-Means Cluster/Hierarchical Cluster/Two-step Cluster
```

**Stata code:**

```
cluster complete v1 v2 v3, measure(absolute)
```

**R code:**

```
d <- dist(data, method = "euclidean")  
fit <- hclust(d, method="ward")
```

Please remember that these codes can receive various specifications depending on the researcher's analysis needs.

## 4.2 Empirical Relations Analysis

Once the indicators are chosen based on the theoretical concepts, the next step is to verify if these indicators have any relation to each other, be it bivariate or multivariate (Babbie, 2010). The difference from this step to the previous one is that when selecting the items we are concerned with their individual relevance, and here we analyze the general framework of indicators and their correlation.

Pearson's correlation indicates an association between variables, that is, a shared variation of the scores of these variables (Figueiredo Filho and Silva Jr., 2009). When two or more indicators show predictability among themselves, that is, knowing the response of one can anticipate the response of the other, there is a relationship between them.

The correlation matrix between the variables / indicators can be executed through the codes:

### **SPSS code:**

Menu Analyze > Correlate > Bivariate > Variables: v1+v2+v3

### **Stata code:**

```
correlate v1 v2 v3
```

### **R code:**

```
cormat <- cor(dados)  
round(cormat, 2)
```

A correlation between indicators may denote that they belong to the same dimension, making it easier to aggregate the data into factors or components (see topic 3 below). Correlation analysis may also indicate which variables will be redundant in the model (Sullivan, 1974).

It may be tempting for the researcher to include several indicators in the index, and thus make it difficult for the individual to eliminate some of them. Therefore, it is always good to keep in mind that these indicators, in the pure semantics of the word, only indicate the concept in a quantified format, not replacing it (Jannuzzi, 2002). Knowing that the concept will have a mathematical representation in the indicator, we should not overestimate it in the construction of the index.

## 4.3 Index Scoring

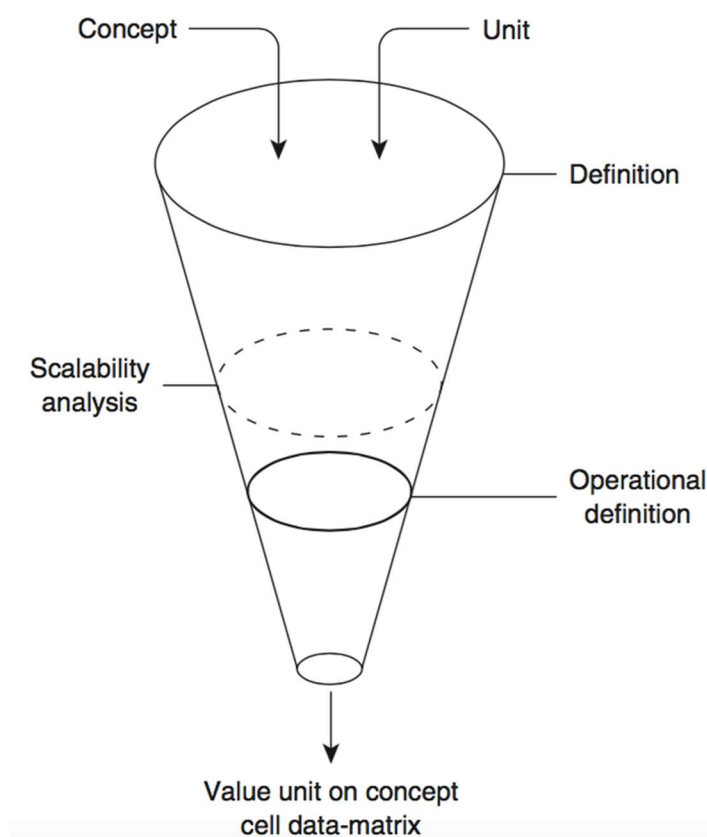
The reader will soon find out that there is no middle ground for this step, or it is either a fun job for the researcher, or a completely unbearable one. Why that? To assign concrete figures to political and sociological concepts is quite arbitrary. Even pain, an example used at the beginning

of the article, is not so palpable to value. The operationalization consists of transforming theoretical concepts into observed variables, either directly or indirectly.

It is important to be able to differentiate the concept from its measurement. A concept can be continuous, that is, it varies in level, and the researcher can still adopt a dichotomous measurement for him (Bollen, 1990). For the case of democracy, a dichotomous measurement would have only the options of values "democratic" and "non-democratic". An index helps you measure the concept continuously.

When scoring the index, we are operationalizing the concept. This is the highest moment of measurement. Pennings, Keman and Kleinnijenhuis (2006) called this procedure a "funnel of operationalization". In the wide mouth of the funnel, which is the definition, we put the concept and the unit. They slip by the first narrowing channel, which is the scalability analysis, then by the second narrowing channel, which is the operational definition, until it reaches the minimum funnel nozzle whose result is the unit value in an array of conceptual cells.

**Figure 3 - "funnel of operationalization"**



Source: Pennings, Keman e Kleinnijenhuis (2006)



There are two potential problems that arise in assigning values. One of them comes when we apply these values to cases. The second problem is in relation to the weight given to the indicators. See below:

**Problem 1:** How to classify them? Why is a given case worth 1 or 2 when it could have 1.5? As for this there is no fixed rule and it is up to the researcher to justify their choice, conscious that you will inevitably be criticized.

**Solution:** Factor analysis is a statistical technique of data reduction and identification of their structure through the inter-correlations of variables (Hair et al, 2005). This process is done by calculating scores that replace the original values of the variables (Hair et al, 2005), as well as the attribution mentioned above, but the factorial analysis does it based on the actual structure of the data, not with arbitrariness.

For applied examples of factorial analysis, and how to select the number of factors, there are the works by Bezerra and Corrar (2006) and Figueiredo Filho et al (2014). The initial basic codes for using this technique are as follows:

**SPSS code:**

Menu Analyze > Dimension Reduction > Factor

**Stata code:**

factor v1 v2 v3 or factor v1-v3

**R code:**

```
fit <- factanal(dados, factors=x, rotation="varimax")
print(fit)
```

**Problem 2:** To illustrate, the researcher can decide that an indicator classifies their cases between 0 and 1, while another indicator does it between 0 and 2. This means giving different weights to the indicators.

**Theoretical solution:** Babbie (2010) argues that while this decision is up to the researcher, the best strategy is to equally weigh the indicators.

**Empirical Solution:** There are some statistical strategies to evaluate how to weigh variables, and even exclude or aggregate some of the index. The Principal Component Analysis, a progression of factorial analysis, allows the grouping of information into factors with minimal loss of original data, and also allows to identify which of these factors contribute more to explain its variance (Hair et al, 2005).

Figueiredo Filho et al (2013) bring the use of principal components analysis in the construction of social indicators. Mendes (2014) applies the technique to his index of institutional development of municipalities. Finally, a classic application is Putnam's (2000) in its social capital index.

To perform a basic principal component analysis, you can use the following codes:

**SPSS code:**

```
Menu Analyze > Dimension Reduction > Factor > Extraction >  
Method:Principal Components
```

**Stata code:**

```
pca v1 v2 v3
```

**R code:**

```
dados.pca <- prcomp(data, center = TRUE, scale. = TRUE)
```

#### 4.4 Index Validation

The last step is index validation - to show that it actually measures what you intend to measure. I have already discussed here what is meant by validity and its role in constructing an index. This is different from the measurement validation process, and the terms should not be confused (Adcock and Collier, 2001). Basically, there are two checks to be done, at micro level and macro level.

At the micro level, we have to check whether an indicator is contributing to the index individually or just reproducing the information from another indicator present (Babbie, 2010). Conversely, an indicator may simply not fully reflect the desired concept. It is necessary to strike a balance between the parsimony proposed by the first situation and the inclusion of components that may arise from the second situation (Adcock and Collier, 2001).

At the macro level, we know that when a study can be generalized beyond the observed cases, it has external validity (De Vaus, 2001). The process of external validation of an index happens when other indexes - thus, different measurements - arrive at the same conclusions as yours (Babbie, 2010). For example, there are several indexes for democracy, which is great for science, but if two indexes conclude divergently whether a country is democratic or not, this becomes problematic.

Another case is the assessment of student learning. In 2015, the IDEB put Pernambuco first in the performance of public high school students. In PISA 2015 (International Student Assessment Program), Pernambuco is in the nineteenth place among the Brazilian states. Where is the problem?

In the selection of indicators, in their measurement, in the score of items? Due to the divergence, there is a possibility that one of the indexes is not valid.

In order to work with this issue, a quali-quantitative analysis can be done between existing indexes. A priori a qualitative analysis goes deep in the construction of each index (selected for analysis - it is important that they have the same cases) to verify if they measure the same things.

The quantitative analysis involves some statistical tests, such as the aforementioned Principal Components Analysis, as well as a correlation between index pairs, as performed by Lorenz (2010) when comparing different indices of constitutional rigidity. A technique that measures the reliability of the index is Cronbach's alpha, through the consistency of the whole scale, the limit for exploratory searches being 0.60 (Hair et al, 2005).

To calculate Cronbach's alpha:

**SPSS code:**

Menu Analyze > Scale > Reliability Analysis > Items: v1 v2 v3 > Model: Alpha

**Stata code:**

.alpha v1-v3, item

**R code:**

```
library(psych)  
psych::alpha(data)
```

The convergence of empirical associations between indicators in general is expected, and when there is divergence the researcher can return to the concept to verify if: its systematization was done incorrectly, or if only the systematized concept is different from that indicator that was sought association (Adcock and Collier, 2001). An outlier may also be upsetting the analysis: scan your data well!

#### **4.5 Other Issues Of Analysis**

Finally, some points to be considered during the construction of indexes. One of them is a recurring problem, the nonexistent data. In the specific literature the reader will find this term with the denomination of missing data. Blank observations in the matrix are common, either systematically for certain variables or cases, when the data does not exist, or randomly, when the data is lost, or it is no longer filled in (in case of a questionnaire for example).

Qualitatively some choices can be made. For a few cases with missing data, they can be excluded from the analysis. You can also choose to fill in one of the available answers based on the others - this is common in questionnaires - or with the middle ground in case of punctuation. Finally, it is also possible to simply interpret what non-existence of data can mean and include it in the analysis (Babbie, 2010).

Quantitatively, once again using statistics, one way out is to check if the indicator that has missing data correlates with another indicator and use that to predict incomplete values and fill them (Little and Rubin, 1987). Care must be taken not to fall into the illusion that these values were not created artificially. A more refined form of this tactic is the Buck method. When faced with a normal multivariate distribution in a covariance matrix, it estimates the regression coefficients of the incomplete variables from the function generated by the complete cases. All this creates some level of distortion, and there is no method that works perfectly for this type of case (Little and Rubin, 1987).

More recent attempts to improve this process come in the technique of multiple imputation, based on Little and Rubin (1987) and previous works. The values are imputed in the empty cells through a parametric model that assumes that the complete data have a joint probabilistic distribution (Lall, 2016).

The use of this technique considers the problems that arise from the simple exclusion of non-existent data, especially in Political Science, where this absence may be non-random, caused by economic development or political institutions (Lall, 2016).

A second point goes beyond the construction of the index to the selection of cases that will receive index treatment. Lorenz (2010) asks: how to choose the countries? After all, when comparing indexes, we have to use the same set of cases. This raises concern about the selection bias. When starting to analyze a biased index, an indicator may show a causal effect that does not exist. One must avoid common errors that lead to bias: to select by the dependent variable, to select according to a fixed point in time, and to select based on path-dependence (Geddes, 1990).

Finally, increasingly sophisticated techniques for the construction of indices have been developed. Based on the example of social development indexes, Foe and Tanner (2012) propose to use Matching Percentiles to aggregate variables, with the advantages that this method deals better with non-existent values, with the processing of binary variables (yes and no, certain and wrong, present and absent), with few countries in the set of cases, and also it doesn't need to assume a linear distribution of the indicators.

A recent venture is also quite interesting: Souza, Graça e Silva (2017) use network analysis of federal deputies on Twitter and Bayesian inference to estimate ideal points (scores) on the scale of political ideology.

This topic serves as a reminder to the reader that despite the elementary criteria, we must be attentive to the various details to be examined. In addition, scientific methods are always evolving in search of better and more reliable results.

## 5. FINAL CONSIDERATIONS

The reader opened this article concerned with the use of indexes in the social sciences. It is a relevant concern. Here I have tried to show that, in science, the anxieties must be solved with rigorous criteria, leaving visible the existent failures and working to mitigate their effects. This is not different in the social sciences, despite the complexity of their objects.

The indexes are often unavoidable, because of the multidimensionality I mentioned at the beginning. Therefore, improving your construction, as I suggested in the last section, is essential. There is no single correct method, and we must take advantage of divergences among researchers to leverage discussions in the social sciences. This article sought to bring different perspectives and suggestions on the path to be followed.

It is important to note that difficulties encountered by researchers and even students are not random. For example, in the paper by Souza, Graça e Silva (2017) the authors comment on the use of high data processing computers in similar works that they try to replicate. They emphasize that in the social sciences, and especially in Brazil, this type of computer is not common access, resulting in greater use of time to perform the same tasks.

Such questions are not trivial, nor easy to solve. Most are not even within reach of the researcher. Despite the limited resources, investing in methodological learning is an important ammunition for the development of science. Access to data and technologies may be comparatively scarce, but the work produced will be rigorous and valid in the construction of scientific knowledge. Methodological articles of all kinds are welcome in Political Science to fit this goal.

As I have been insisting since the beginning of this paper, the use of research techniques should always be accompanied by good notions of measurement theory. It is not by chance that the first step in index construction lies in the empirical delimitation of the concept. Complex concepts are great for theoretical debate, but without adequate measurement, the empirical development of the discipline is lost. This recommendation is not only for students who study how to create indexes but also for teachers to include measurement theory in their methodology courses.

What themes and concepts do you, reader, think are still under-explored in Political Science or another area of knowledge? Is it possible that this introduction to the use of indexes has helped you, or even encouraged, to fill that gap? Or do you think that some existing measurement is not being done correctly, and you set out to replicate using more rigorous methods? If the answers are positive, the article has fulfilled its role.

## REFERENCES

- ADCOCK, Robert and COLLIER, David. (2001) "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95:3 (September), 529-46.
- BABBIE, Earl. (2010) "Indexes, Scales and Typologies." In: *The Practice of Social Research*. Twelfth Edition. Wadsworth, Cengage Learning.
- BEZERRA, Francisco Antonio, and Luiz J. CORRAR. (2006). "Utilização da análise fatorial na identificação dos principais indicadores para avaliação do desempenho financeiro: uma aplicação nas empresas de seguros." *Revista Contabilidade & Finanças* 17, no. 42: 50-62.
- BOLLEN, Kenneth A. (1990). "Political democracy: Conceptual and measurement traps." *Studies in Comparative International Development* 25, no. 1: 7-24.
- BRADY, Henry and COLLIER, David. (2004) *Rethinking Social Inquiry*. Rowman and Littlefield.
- CUKIERMAN, Alex; WEBB, Steven B.; NEYAPTI, Bilin. (1992) "Measuring the Independence of Central Banks and Its Effect on Policy Outcomes." *The World Bank Economic Review*. Vol. 6, No. 3: 353-398.
- ESI Triage Research Team (2004). Disponível em: <<http://www.esitriage.org/index.asp>>
- DAHL, Robert. (1997) *Poliarquia: participação e oposição*. 1ed. 3 reimp. Coleção Clássicos, EDUSP. São Paulo.
- DEMOCRACY BAROMETER (2016). Disponível em: <[www.democracybarometer.org](http://www.democracybarometer.org)>
- DE VAUS, David (2001). "Tools for research design." In *Research Design in Social Research*, ed. David De Vaus. Thousand Oaks, CA: Sage.
- FIGUEIREDO FILHO, D. B. ; SILVA JÚNIOR, J. A. (2009). "Desvendando os Mistérios do Coeficiente de Pearson (r)." *Política Hoje* (UFPE. Impresso), v. 18, p. 115-146.
- FIGUEIREDO FILHO, Dalson Britto, Enivaldo Carvalho da ROCHA, Ranulfo PARANHOS, Anderson Henrique SILVA, José Alexandre SILVA JR, Lucas SILVA, and Dáfni Priscila ALVES. (2014). "Análise Fatorial Garantida ou o Seu Dinheiro de Volta: Uma Introdução à Redução de Dados." *Revista Eletrônica de Ciência Política* 5, no. 2.
- FIGUEIREDO FILHO, Dalson Britto, Ranulfo PARANHOS, Enivaldo Carvalho da ROCHA, and Romero G. MAIA. (2013). "Análise de componentes principais para construção de indicadores sociais." *Rev. Bras. Biom* 31, no. 1: 61-78.
- FOE, R, & TANNER, J.C. (No. 2012-04) "Methodology of the Indices of Social Development". *ISD Working Paper Series*. Retrieved from <<http://hdl.handle.net/1765/50510>>
- FREEDOM HOUSE (2016). Disponível em: <<https://freedomhouse.org>>

GEDDES, Barbara. (1990) "How the cases you choose affect the answers you get: selection bias in comparative politics". *Political Analysis*. 2 (1): 131-150.

GILL, J. (2006). *Essential mathematics for political and social research* (pp. 271-277). Cambridge: Cambridge University Press.

HAIR Jr., Joseph F.; BLACK, William C.; BABIN, Barry J.; ANDERSON, Rolph E. (2005) *Análise Multivariada de Dados*. Quinta edição. Bookman, Porto Alegre.

HANSEN, Jaqueline Resmini. (2016) "*No computador, na rua ou no smartphone: condicionantes ou covariantes do engajamento online*." PhD diss., Dissertação de mestrado, Programa de Pós-Graduação em Ciências Sociais, UEM, Maringá.

JANNUZZI, Paulo de Martino. (2002) "Considerações sobre o uso, mau uso e abuso dos indicadores sociais na formulação e avaliação de políticas públicas municipais." *RAP*, 36(1):51-72. Rio de Janeiro.

KELLSTEDT, Paul M; WHITTEN, Guy D. (2009) *The Fundamentals of Political Science Research*. Cambridge University Press.

KING, Gary; KEOHANE, Robert and VERBA, Sidney. (1994) *Designing Social Inquiry: scientific inference in qualitative research*. Princeton University Press.

KUBRUSLY, Lucia Silva. (2001) "Um procedimento para calcular índices a partir de uma base de dados multivariados." *Pesquisa Operacional* 21, no. 1: 107-117.

LALL, Ranjit. (2016). "How multiple imputation makes a difference." *Political Analysis* 24, no. 4: 414-433.

LANDMAN, Todd. (2008) *Issues and Methods in Comparative Politics: an introduction*. Third Edition. Routledge, Taylor & Francis Group e-Library.

LIJPHART, Arend. (2008) *Modelos de democracia: desempenho e padrões de governo em 36 países*. 2a. ed. Rio de Janeiro: Civilização Brasileira.

LITTLE, Roderick and RUBIN, Donald B. (1987) *Statistical Analysis with Missing Data*. New York: John Wiley & Sons.

LORENZ, Astrid. (2005) "How to Measure Constitutional Rigidity: Four Concepts and Two Alternatives." *Journal of Theoretical Politics*, 17(3): 339–361. SAGE.

MAHONEY, James e VANDERPOEL, Rachel Sweet. (2014) "Set Diagrams and Qualitative Research." *Comparative Political Studies*, 2015 48: 65. SAGE Publications.

MAINWARING, Scott, Daniel BRINKS, and Aníbal PÉREZ-LIÑÁN. (2001) "Classifying political regimes in Latin America, 1945-1999." *Dados* 44, no. 4: 645-687.

MENDES, Renata. (2014) "Brazilian National Housing Policy: Institutional capacity and housing deficit." *Politikon: IAPSS Political Science Journal*, Vol. 23, June.



MERKEL, Wolfgang and BOCHSLER, Daniel (project leaders) (2016); Bousbah, Karima; Bühlmann, Marc; Giebler, Heiko; Hänni, Miriam; Heyne, Lea; Müller, Lisa; Ruth, Saskia; Wessels, Bernhard. *Democracy Barometer. Methodology. Version 5*. Aarau: Zentrum für Demokratie.

MOORE, W. H.,; SIEGEL, D. A. (2013). “*A mathematics course for political and social research*”. Princeton University Press.

MOZZAFFAR, S., & SCHEDLER, A. (2002) “The Comparative Study of Electoral Governance—Introduction.” *International Political Science Review*, 23, 5–27.

NUNNALLY, Jum C., and Ira H. BERNSTEIN. (1978) “*Psychological theory*.” New York, NY: MacGraw-Hill.

PENNINGS, Paul; KEMAN, Hans; KLEINNIJENHUIS, Jan. (2006) *Doing Research in Political Science*. 2nd edition. SAGE Publications. London.

POGREBINSCHI, Thamy. (2010) “Democracia Pragmática: Pressupostos de uma Teoria Normativa Empiricamente Orientada.” *DADOS – Revista de Ciências Sociais*, Rio de Janeiro, vol. 53, n 3, pp. 657 a 693.

PUTNAM, Robert D. (2000). *Bowling alone: America’s declining social capital*. Simon & Schuster, New York.

ROBERTS, Fred S. (2009) *Measurement Theory*. Cambridge University Press (1985), digitally printed version.

SANTOS, B. R. L., TEODORO, J. V., & FACCIN, A. C. T. M. (2015). “Análise de componentes principais dos indicadores de qualidade de vida”. *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, 3(1).

SARTORI, Giovanni. (1970) “Concept Misinformation in Comparative Politics.” *American Political Science Review*. 64:4 (December), 1033-53.

SOUZA, Rafael Martins de, Luís Felipe Guedes da GRAÇA, and Ralph dos Santos SILVA. (2017) “Politics on the Web: Using Twitter to Estimate the Ideological Positions of Brazilian Representatives.” *Brazilian Political Science Review* 11, no. 3.

STRAUSS, Leo. (2009) *Direito Natural e História*. - Reimp. (Biblioteca de teoria política; 3). Edições 70. Lisboa.

SULLIVAN, John L. (1974). “Multiple indicators: Some criteria of selection.” In *Measurement in the social sciences* / Hubert Blalock (ed.), pp. 243-269. The Macmillan Press.

TAROUCO, Gabriela. (2012) “Governança eleitoral: incentivos institucionais e competição partidária”. *8o Encontro da Associação Brasileira de Ciência Política*. Gramado.

VAN EVERA, Stephen. (1997) “Hypotheses, laws and theories: a user’s guide.” In Stephen Van Evera. *Guide to methods for students of Political Science*. Cornell University Press.

WEBER, Max. (1982) “A Ciência como Vocação”. In: *Ensaio de Sociologia*. LTC - Livros Técnicos e Científicos Editora S.A.