

**ASSOCIAÇÃO LATINOAMERICANA DE CIÊNCIA POLÍTICA**  
**XVIII CONGRESSO LATINO-AMERICANO DE CIÊNCIA POLÍTICA**  
**PONTIFICA UNIVERSIDAD CATÓLICA DEL PERÚ**  
**LIMA – PERÚ**

**“ÍNDICE BRASILEIRO DE CRIMINALIDADE”**

**AUTORES**

**Jose Alexandre da Silva Júnior (UFAL)**  
[jasjunior2007@yahoo.com.br](mailto:jasjunior2007@yahoo.com.br)

**Ranulfo Paranhos (UFAL)**  
[ranulfoparanhos@me.com](mailto:ranulfoparanhos@me.com)

**Dalson Britto Figueiredo Filho (UFPE)**  
[dalsonbritto@yahoo.com.br](mailto:dalsonbritto@yahoo.com.br)

**Enivaldo Carvalho da Rocha (UFPE)**  
[eni-rocha@hotmail.com](mailto:eni-rocha@hotmail.com)

**Willber da Silva Nascimento (UFPE)**  
[nascimentowillber@gmail.com](mailto:nascimentowillber@gmail.com)

## Indicador Brasileiro de Criminalidade

“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.”

(ROBERTS, F. S., 1985, p. 01)

Como mensurar a criminalidade no Brasil? Comumente, são utilizadas duas estratégias: 1) escolha de uma *proxy* (número de homicídio) e 2) análise de um conjunto de dados. Para literatura especializada, ambas tendem a gerar medidas de baixa validade e confiabilidade. Esse trabalho tenta corrigir esses erros e propõe a criação de um indicador de criminalidade. Precisamente, utilizamos a análise fatorial para reduzir dez variáveis: estupro, tentativa de estupro, lesão corporal, homicídio, tentativa de homicídio, roubo (veículo, instituições financeiras e outros), latrocínio, porte ilegal de armas, porte de entorpecentes e tráfico de entorpecentes. Os dados são referentes às ocorrências registradas entre 2010 e 2012 nos 26 estados mais o Distrito Federal. Os resultados sugerem a construção de um indicador com três componentes ortogonais, capaz de explicar 70% da variância total.

Palavras-chaves – Mensuração; Criminalidade; Análise Fatorial; Homicídio;

### INTRODUÇÃO<sup>1</sup>

A escalada da violência é um dos fenômenos mais alardeado pela mídia brasileira. O assunto também desperta cada vez mais interesse na academia. Entretanto, esse destaque não foi acompanhado por um esforço para mensuração do fenômeno. Em geral, a ocorrência de Crimes Violentos Letais e Intencionais (CVLI's) tem sido utilizada como indicador da criminalidade no Brasil (BEATO, PEIXOTO e ANDRADE, 2004). São raros os trabalhos que procuram representar o fenômeno a partir de uma perspectiva multivariada. Salvo engano, não há trabalhos que analise o compartilhamento de variância entre as múltiplas modalidades de crimes.

De acordo com a literatura, as consequências dessa lacuna podem ser devastadoras para análise do fenômeno. Tanto a escolha de uma *proxy* (como número de homicídio) quanto a análise de um conjunto de dados tratados isoladamente podem gerar vieses sistemáticos de mensuração (BLALOCK, 1979; JACOBSON E LALU,

---

<sup>1</sup> Agradecemos ao pesquisador Denisson Silva por comentários e coletas de dados. Eventuais imprecisões são exclusivamente creditadas aos autores. Nosso trabalho é financiado por duas principais fontes: CNPQ e FAPEAL.

1974). Em outras palavras, ambas tendem a criar medidas com baixa validade e confiabilidade (ZELLER E CARMINES, 1980). Curiosamente, a literatura especializada nacional não procurou investigar o impacto dessa tendência<sup>2</sup>. Nesse trabalho, seguimos a literatura e apostamos na multidimensional da criminalidade. Nossos objetivos serão: 1) analisar a estrutura de correlação entre diferentes modalidades de crime; 2) verificar a possibilidade de reduzi-las e 3) construir um indicador de criminalidade. Para tanto, utilizamos a análise fatorial de componentes principais e observamos dez modalidades de crime: estupro, tentativa de estupro, lesão corporal, homicídio, tentativa de homicídio, roubo (veículo, instituições financeiras e outros), latrocínio, porte ilegal de armas, porte de entorpecentes e tráfico de entorpecentes. Em resumo, esse artigo propõe uma investigação criteriosa da covariação dos dados sobre crimes no Brasil. Os dados são referentes às ocorrências registradas entre 2010 e 2012 nos 26 estados mais o Distrito Federal.

O artigo está organizado em quatro seções. Na primeira, esboçamos uma breve revisão teórica sobre a multidimensionalidade do crime. Na sequência, abordamos as vantagens e limitações das diferentes estratégias de mensuração de fenômenos não diretamente observáveis. Na terceira seção, apresentamos nossas hipóteses e os principais resultados da análise. Na quarta e última, tecemos as considerações finais.

## 1. MULTIDIMENSIONALIDADE DO CRIME

Becker (1968), Stigler (1970) e Ehrlich (1973) são os “clássicos” que investigaram as múltiplas modalidades de crimes. O foco das investigações concentrou-se no custo social da transgressão. Becker (1968) estava interessado em descobrir as razões para a variação das punições. De acordo com ele, “o crime é uma importante atividade econômica ou uma “indústria”” (p.170)<sup>3</sup>. Mais que isso, ele acreditava ser possível calcular o custo social de cada atividade criminosa. Elementos como o prejuízo gerado, custo de repressão e as oportunidades para o crime ajudam a explicar a variação da punição. Para Becker (1968) a punição é uma função do esforço social para combater a atividade criminosa. Da mesma forma, Stigler (1970) acredita que “there is one

---

<sup>2</sup> Antes da realização desse trabalho foi realizada uma busca sistemática por trabalhos que discutem a mensuração da criminalidade no Brasil. Pesquisou-se pelos termos “criminalidade”, “medidas de criminalidade”, “mensuração de criminalidade” no portal *Scielo* e nos *sites* de busca Google e do Fórum Brasileiro de Segurança Pública.

<sup>3</sup> No original: “crime” is an economically important activity or “industry”.

decisive reason why the society must forego “complete” enforcement of the rule: enforcement is costly” (STIGLER, 1970, p. 56). Isso significa que parte da variação da ocorrência de um crime deve-se ao custo que a sociedade está disposta a pagar para combatê-lo. Esse esforço altera principalmente o custo e o risco envolvido na atividade criminosa. Ou seja, a sociedade paga um preço para tornar o crime menos atraente <sup>4</sup>. Adicionalmente, Ehrlich (1973) defende que as atividades legais também são parte da equação. Mais precisamente, os ganhos e os custos de atividades legais e ilegais são elementos indispensáveis para análise da “economia do crime”. A ocorrência das atividades criminosas varia segundo esses quatro elementos (EHLICH, 1973). Em outros termos, ela é uma função da estrutura de oportunidades (custos, riscos e ganhos) associada ao crime (HEINEKE, 1978; WOLPIN, 1978; BLOCK e HEINEKE, 1975; EIDE, 1999; SANDMO, 1972; KOLM, 1973; SINGH, 1973).

Segundo Ehrlich (1973), as diferentes modalidades de crime é um fator da variação dessa estrutura de oportunidade. Para Eide (1990), “the kinds of gains obtained from a criminal act vary, depending on the type of crime and the individual criminal” (EIDE, 1990 p. 351). Ehrlich (1973), Entorf e Spengler (1998) e Gould, Weinberg e Mustard (2002) argumentam que uma medida básica é agrupar os crimes em: contra a vida e contra a pessoa. Para Ehrlich (1973), “crimes against the person may be motivated primarily by hate or passion: phenomena involving interdependencies in utilities among individuals” (p.532). Em geral, os crimes contra a pessoa ocorrem entre indivíduos com contato próximo e frequente (EHLICH, 1973; ENTORF e SPENGLER, 1998). Segundo Entorf e Spengler (1998), o “mercado do crime” tem dinâmica distinta para crimes contra a pessoa e crimes contra o patrimônio. Mais que isso, eles mostram que as variáveis que explicam a frequência de um tipo, não explicam a ocorrência dos demais (ENTORF e SPENGLER, 1998). Gould, Weinberg e Mustard (2002) evidenciam que a curva de crimes contra a propriedade e violentos apresentam dinâmicas distintas nos estados unidos no período entre 1979 e 1997. Eles mostram ainda que as condições do mercado de trabalho (desemprego e escolaridade) têm impacto distinto nas duas modalidades de crime.

Em resumo, acredita-se que o esforço da sociedade para o combate ao crime não é igualmente distribuído entre as modalidades (BECKER, 1968; STIGLER, 1970;

---

<sup>4</sup> Para os teóricos da escolha racional a decisão de cometer um crime é resultado de um processo de maximização de utilidade esperada. Entram na equação o ganho, o risco e o custo da ação. Por eles, o indivíduo decide se comete ou não o ato criminoso (BECKER, 1968; GOTTFREDSON e HISCHI, 1990).

EHRLICH, 1973 e PETER e TAYLOR, 1977). Em geral, a sociedade está disposta a gastar mais com os crimes que geram maior perda coletiva. Na prática, isso significa que algumas modalidades de crime podem se tornar mais atraente que outras. Em tese, crimes contra a vida e crimes contra o patrimônio tem mercado distintos. Isso justifica a análise separada para cada categoria (ENRLICH, 1973; ENTORF e SPENGLER, 1998 e GOULD, WEINBERG e MUSTARD, 2002).

## 2. ESTRATÉGIAS DE MENSURAÇÃO

(linkar com a teoria) Três estratégias são comumente utilizadas para mensurar um fenômeno abstrato: 1) escolher um indicador; 2) construir um índice ou 3) elaborar uma medida com múltiplos indicadores. Segundo Coombs (1964), **cada estratégia** corresponde a um conjunto específico de regras que relaciona o dado à teoria. **Cada estratégia** se aplica a um fenômeno específico e possui vantagens e desvantagens. O quadro abaixo ajuda a compará-las:

**Quadro 01 – Comparação das Estratégias de Mensuração**

Estratégia	Fenômeno	Vantagem	Desvantagem
Indicador Único	Simple (univariado) e unidimensional.	Inteligível e mais simples de reproduzir.	Mais ausência de dados; Sujeito a erros não especificados e a baixa validade.
Índice	Complexo (multivariado) e unidimensional.	Representa melhor o conceito, menos sujeito a erros sistemáticos.	Sujeito a Heterocedasticidade e a estrutura de correlação ambígua.
Múltiplos Indicadores	Complexo (multivariado) e multidimensional.	Mais criteriosa na ligação entre conceito e medida. Menos sujeito a erros sistemáticos e ambigüidades	Precisa de teoria sobre “causa e efeito” (variáveis e conceito). Requer hipótese sobre a estrutura de correlação entre as variáveis. Necessita de conhecimento sobre o método de agregação das variáveis.

Fonte: Elaboração própria baseado em JACOBSON e LALU (1974).

Ao optar por um indicador o pesquisador assume que o fenômeno é univariado. Ou melhor, ele assume que o fenômeno pode ser representado/dimensionado por apenas uma variável observada. Isso implica caracterizar o fenômeno como simples e unidimensional. Para Conster (1969) e Heise (1969) a grande vantagem dessa estratégia está na simplicidade da medida. Tende a facilitar a interpretação e a replicação. Todavia, corre-se o risco de ignorar a multidimensionalidade dos fenômenos e, conseqüentemente, elaborar uma medida contaminada por erros sistemáticos.

No caso da construção de um índice o intuito é mensurar fenômenos multivariados. Para Lazarsfeld e Rosenberg (1955) os índices servem como medidas sínteses que identificam um conjunto de atributos que compõem um conceito. Existem várias formas de executar essa estratégia. Elas variam segundo: 1) número de itens; 2) distribuição dos pesos e 3) forma como os itens são combinados (JACOBSON e LALU, 1974). De acordo com Hauser e Goldberger (1971), a estratégia mais comum é a mera adição dos itens sem atribuição de pesos. O índice é adequado para dimensionar fenômenos multivariados cuja composição é clara e conhecida. Isso porque, costuma prescindir de teorias sobre a relação entre os itens, ignorar correlações ambíguas entre eles e desconsiderar a correlação entre os erros devido a variáveis omitidas no modelo (SULLIVAN, 1974).

Por fim, a medida elaborada a partir de múltiplos indicadores é adequada para fenômenos multivariados e multidimensionais (BLALOCK, 1967; COSTNER, 1969; JACOBSON e LALU, 1974; SULLIVAN, 1974; ASHER, 1993). Diferentes das demais, essa estratégia permite observar a relação de ‘causa e efeito’ entre cada variável observada e o conceito. Com ela, pode-se calcular o erro de estimação da medida e identificar diferentes tipos de vieses (COSTNER, 1969). Portanto, a estratégia exige: 1) explicações teóricas sobre a relação entre as variáveis observáveis; 2) explicações teóricas sobre a relação entre estas variáveis e o conceito (BLALOCK, 1971) e 3) conhecimento sobre os diferentes métodos de estimação da medida (HAUSER e GOLDBERGER, 1971).

### **2.3 MENSURANDO CRIMINALIDADE**

No Brasil, a maioria dos trabalhos utiliza a taxa de homicídios como indicador de criminalidade. Portanto, implícita ou explicitamente assume que a criminalidade é um fenômeno simples e unidimensional. Ou melhor, parte do pressuposto de que crimes

como estupro, roubo, lesão corporal, dentre outros, têm um forte compartilhamento de variância com homicídio<sup>5</sup>. Contrariamente, a literatura especializada aponta um bidimensionalidade do fenômeno. Indica que crimes contra a pessoa e crimes contra o patrimônio tem dinâmicas específicas. Isso está refletido na gradação de esforços que a sociedade faz para reprimir os crimes. Nesse trabalho, seguimos a literatura e acreditamos que a criminalidade é fenômeno multidimensional. Por isso, optamos por mensurar a criminalidade a partir de múltiplos indicadores. Portanto, nos propomos a: 1) analisar a estrutura de correlação entre diferentes modalidades de crime; 2) verificar a possibilidade de reduzi-las e 3) construir um indicador de criminalidade.

Para tanto, utilizamos a Análise Fatorial de Componentes Principais (AFCP). De acordo com Hair et al (2005) “factor analysis is an interdependence technique whose primary purpose is to define the underlying structure among the variables in the analysis” (HAIR ET AL, 2005 p. 94). Essa técnica serve principalmente para “to summarize patterns of correlation among observed variables, to reduce a large number of observed variables to a smaller number of factors (...) or to test a theory about the nature of underlying processes” (TABACHNICK e FIDELL, 2012 p.95). Portanto, a técnica tem três funções básicas: 1) sumarização, 2) redução e 3) validação dos dados.

### *Sumarização dos Dados*

Para Hair et al (2005) “the goal of data summarization is achieved by defining a small numbers of factors that adequately represent the original set of variable” (HAIR ET AL, 2005 p. 97). Portanto, o foco é a correlação entre as variáveis que se pretende sumarizar. A sumarização dos dados depende do método de extração e de rotação dos fatores. Existem vários métodos de extração, para analisar a criminalidade utilizaremos os componentes principais<sup>6</sup>. Para Tabachnick e Fidell (2012) a análise fatorial por componentes principais busca maximizar a variância extraída por componentes ortogonais. O método é adequado para observar a quantidade de fatores necessários para mensurar um fenômeno abstrato. Toda variância é distribuída entre os componentes,

---

<sup>55</sup> A magnitude dessa variância pode ser dimensionada pela correlação existente entre as modalidades de crime. Para ser considerada forte o coeficiente da correlação precisa estar acima de 0.8 (ver DANCEY e REIDY, 2006).

<sup>6</sup> O Pacote estatístico SPSS v. 19 – oferece seis métodos de extração do fator: principal components (PCA), principal factors, maximum likelihood factoring, image factoring, alpha factoring, unweighted e generalized (weighted) least squares factoring. A principal diferença entre eles é o critério que cada um utiliza para maximizar a variância ou minimizar a correlação dos resíduos (TABACHNICK e FIDELL, 2012).

incluindo o erro e variância específica de cada variável (TABACHNICK e FIDELL, 2012 p.640) <sup>7</sup>. Em resumo, na etapa de extração verificamos a dimensionalidade do indicador de criminalidade.

O resultado da extração está diretamente ligado à correlação entre as variáveis observadas. Padrões de correlação claros facilitam a decisão sobre o número de fatores a serem extraídos (JACOBSON e LALU, 1974; SULLIVAN, 1974; ASHER, 1993). Quando o padrão de correlação é obscuro, a rotação dos fatores é acionada <sup>8</sup>. A rotação torna mais clara a contribuição de cada variável para construção dos fatores. Ela recalcula as cargas fatoriais carregadas por cada um deles (TABACHNICK e FIDELL, 2012). Existem dois tipos de rotação: ortogonal e oblíquo. A diferença maior entre eles está na correlação entre os fatores extraídos. No ortogonal, os fatores são independentes entre si, no oblíquo eles preservam algum nível de correlação. Para análise da criminalidade iremos observar as soluções oferecidas por duas técnicas, uma de cada tipo. No ortogonal, utilizaremos a *varimax*, cujo objetivo é minimizar a complexidade dos fatores pela maximização da variância das cargas de cada fator (TABACHNICK e FIDELL, 2012). A *varimax* tende a redistribuir proporcionalmente a variância entre os fatores, tornando-os igualmente importantes<sup>9</sup>. A utilização dessa técnica implica duas apostas: 1) a criminalidade é um fenômeno multidimensional e 2) suas dimensões não estão correlacionados entre si. No oblíquo, utilizaremos a *direct oblimin* – cujo objetivo é simplificar os fatores pela minimização dos produtos cruzados das cargas. Essa técnica permite vários níveis de correlação entre os fatores<sup>10</sup>. Portanto, a técnica aposta na dependência entre as dimensões.

---

<sup>7</sup> O método cria tantos componentes quantas variáveis houver. O primeiro fator é resultado do máximo compartilhamento linear entre as variâncias das variáveis observadas. O segundo é formado pela correlação residual, ou seja, “o compartilhamento linear de variância entre as variáveis que extrai o máximo de variância não correlacionada com o primeiro componente” (TABACHNICK e FIDELL, 2012 p.640).

<sup>8</sup> Portanto, o primeiro passo da análise é o exame da matriz de correlação e de covariância entre as variáveis observadas. Tentar encontrar nelas um padrão de correlação a partir do sinal, da magnitude e da significância estatística dos coeficientes (HAIR ET AL, 2005).

<sup>9</sup> Além do *varimax* existem ao menos duas outras técnicas de rotação ortogonal: *quartimax* e *equamax*. A primeira busca minimizar a complexidade das variáveis minimizando a variância das cargas de cada variável. A segunda busca simplificar ambos – fatores e variáveis. Pode ser considerada uma técnica intermediária entre a *varimax* e a *quartimax* (TABACHNICK e FIDELL, 2012).

<sup>10</sup> Além do *direct oblimin* existem ao menos três outras técnicas de rotação oblíqua: *direct quartimin*, *promax* e *procrustes*. A primeira simplifica fatores minimizando a soma do produto cruzado do quadrado das cargas na matriz padrão. Permite apenas altas correlações entre os fatores. A segunda rotaciona fatores ortogonais em posições oblíquas. Por fim, a *procrustes* toma como referência uma matriz alvo de cargas definida pelo pesquisador. Para mais detalhes ver Gorsuch (1983).



### *Redução dos Dados*

Segundo Hair et al (2005) a análise fatorial também pode ser utilizada para “identifying representative variables from much large set of variables for use in subsequent multivariate analyses (HAIR ET AL, 2005 p.98). A meta principal é conservar as características das variáveis originais. Para tanto, três estratégias podem ser utilizadas: 1) adotar uma variável substituta; 2) criar uma variável síntese e 3) utilizar os escores dos fatores (HAIR ET AL, 2005). Para análise da criminalidade nos estados brasileiros comparamos três estratégias distintas.

A escolha de uma variável substituta pode ser teórica e/ou metodologicamente orientada. Em geral, escolhe-se a variável de maior contribuição para construção de cada componente. Dessa forma, reduz-se a arbitrariedade da escolha e aproveitam-se as vantagens de adotar um indicador simples (inteligível e replicável). Todavia, tanto dentro como fora da Academia é comum a escolha arbitrária da variável substituta. Normalmente, opta-se por dimensionar a criminalidade a partir do número de ocorrência de homicídios. Por fim, utilizar os escores dos fatores é opção adequada quando: 1) as comunalidades são altas e 2) a relação entre as variáveis e o fator é clara (GUTTMAN, 1955; HEERMANN, 1963; GORSUCH, 1983). Nessas condições, os escores são uma combinação entre valores de variáveis com altas cargas fatoriais (HAIR ET AL, 2005). Existem ao menos três métodos de composição desses escores: 1) Regressão; 2) Bartlett e 3) Anderson-Rubin. A diferença maior entre eles é a correlação entre os escores gerados e os fatores (GORSUCH, 1993). Nesse texto, utilizaremos um método que gera escores não correlacionados (Anderson-Rubin). O método criado por Anderson-Rubin (1956) produz fatores não correlacionados entre si, mesmo se os fatores estiverem correlacionados (TABACHNICK e FIDELL, 2012). Em resumo, pretendemos contrapor três formas de mensurar a criminalidade: 1) variável representante metodologicamente escolhida; 2) variável arbitrariamente escolhida e 2) composição de fatores.

### *Validação da Medida*

Existem dois parâmetros básicos para avaliar as propriedades de uma medida: confiabilidade e validade. A confiabilidade corresponde à chance de um experimento, teste ou procedimento de mensuração oferecer resultados semelhantes sempre que repetidos (ZELLER E CARMINES, 1979). A validade é a proporção em que a medida mensura aquilo a que ela se propõe, ou seja, é a correspondência entre o conceito e a medida. A ausência de indicadores que contribuem para construção de dois fatores é sinal da confiabilidade da medida (SULLIVAN, 1974; GORDON, 1968).

A literatura especializada recomenda alguns procedimentos para avaliar a confiabilidade e validade das medidas construída por análise fatorial. Precisamente: 1) análise da matriz de correlação residual, 2) fracionamento da amostra e 3) perspectiva confirmatória. Todos serão utilizados na análise da criminalidade dos estados brasileiros. A matriz de correlação residual é uma medida da utilização da correlação entre as variáveis originais para construção dos fatores. Precisamente, ela corresponde a diferença entre a matriz de correlações de variáveis originais e a matriz de correlações reproduzidas<sup>11</sup>. Gordon (1968) indica a replicação da análise com uma amostra diferente ou dividida para verificar se os resultados são robustos. A análise será considerada mais robusta com a permanência do padrão de correlação entre as variáveis, do número de fatores e da distribuição das cargas fatoriais (GORDON, 1968). Por fim, para Hair et al (2005) a perspectiva confirmatória é o método mais direito para validar os resultados de uma análise fatorial. Aqui utilizaremos essa perspectiva a partir do método de máxima verossimilhança. De acordo com Brown (2006) e Thompson (2004), esse método procura estimar parâmetros que maximizem a probabilidade de observar resultados iguais, caso os dados sejam novamente coletados na mesma população (BROWN, 2006)<sup>12</sup>. Para tanto, precisa respeitar alguns pressupostos: 1) amostras grandes (assintóticas); 2) indicadores (variáveis observadas) contínuos e com 3) distribuição normal (BROWN, 2006). Respeitado os pressupostos, a técnica é adequada para testar a confiabilidade dos fatores gerados.

---

<sup>11</sup> Matriz de correlação entre variáveis estimadas a partir das cargas fatoriais extraídas e rotacionadas.

<sup>12</sup> Os principais parâmetros estimados na análise fatorial são: factor loadings, factor variances and covariances, indicator error variances and possibly error covariances (THOMPSON, 2004 p. 72).

## Quadro 02 – Resumo da Metodologia

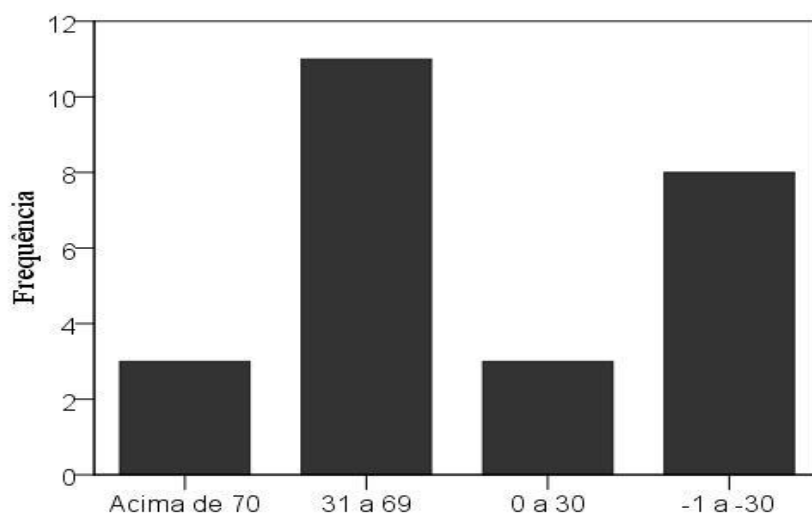
Questão	Como mensurar a criminalidade no Brasil?	
Verificações	Dimensionalidade; Distribuição de Cargas; Generalização.	
População	26 Estados Brasileiros + Distrito Federal	
Período	2010 - 2012	
Variáveis	Estupro, Tentativa de Estupro, Tentativa de Homicídio, Homicídio, Lesão Corporal, Porte Ilegal de Armas, Roubo de Veículo, Roubo a Instituição Financeira, Latrocínio, Tráfico de Entorpecentes, Porte de Entorpecentes.	
Técnica	Análise Fatorial	
	Extração	Componentes Principais.
	Rotação	Varimax e Direct Oblimin.
	Redução	Variável Substituta (escolha metodológica ou arbitrária) e Composição de fatores (Anderson-Rubin).
	Validação	Matriz de correlação residual, fracionamento da amostra e Perspectiva confirmatória ( <i>Generalized Last Square</i> ).

Fonte: Elaboração dos Autores

## RESULTADOS

O primeiro passo da análise é o exame da matriz de correlação e de covariância das variáveis observadas. Tentar encontrar um padrão a partir do sinal, da magnitude e da significância estatística dos coeficientes (HAIR ET AL, 2005). O gráfico abaixo auxilia nessa tarefa.

**Gráfico 01 - Coeficientes de Correlação<sup>13</sup>**



Fonte: Elaboração dos Autores

<sup>13</sup>Análise realizada sem a presença de *outliers* (casos a mais de três desvios padrões da média). Apenas coeficientes com significância estatística foram reportados.

A maior parte dos coeficientes tem magnitude acima de 0,30 (14; 56%), considerando apenas os coeficientes positivos com significância estatística ( $p < 0,05$ ). Esse resultado é animador para a análise fatorial (HARIR ET AL, 2005). Todavia, parte dos coeficientes (8; 32%) tem sinal negativo. Embora preocupante esse resultado não desautoriza a utilização de uma análise fatorial (TABACHNICK e FIDELL, 2012). Para tirar a dúvida, é importante observar se a amostra selecionada é adequada para realização da técnica. Para tanto, a literatura recomenda a apreciação dos resultados do Bartlett Test of Sphercity (BTS) e do Kaiser-Meyer-Olkin (KMO). As tabelas abaixo apresentam os resultados.

**Tabela 01** – Adequação da Amostra

Testes	Componentes Principais	
	KMO	Maginute
BTS	$\chi^2$ Aproximado	305,624
	Grau de Liberdade	55
	p	0,000

Fonte: Elaboração dos Autores

Ambos os testes autorizam a análise. De acordo com a literatura 0,50 é patamar mínimo aceitável para o KMO. Adicionalmente, o BTS precisa apontar que a amostra é adequada ( $p < 0,05$ ). Sendo assim, os resultados autorizam o prosseguimento da análise.

#### *Sumarização dos Dados*

Como visto anteriormente, essa etapa consiste na extração e rotação dos fatores. Portanto, nosso primeiro passo é analisar as comunalidades e a variância explicada pelos fatores com autovalores acima do limite recomendado pela literatura (1,00) (HAIR ET AL, 2005). As tabelas abaixo reportam os resultados:

**Tabela 02 - Comunalidades**

Variáveis	Componentes Principais	
	Inicial	Extração
Estupro	1,000	0,678
Tentativa de Estupro	1,000	0,638
Tentativa de Homicídio	1,000	0,818
Homicídio	1,000	0,835
Lesão Corporal	1,000	0,763
Porte Ilegal de Armas	1,000	0,714
Roubo de Veículo	1,000	<b>0,383</b>
Roubo a Instituição Financeira	1,000	0,725
Latrocínio	1,000	0,729
Tráfico de Entorpecentes	1,000	0,707
Porte de Entorpecentes	1,000	0,711

Fonte: Elaboração dos Autores

**Tabela 03 - Soma do Quadrado das Cargas Extraídas**

Componentes	Total	Variância (%)	Variância Acumulada (%)
1	4,436	40,328	40,328
2	2,246	20,418	60,746
3	1,019	9,266	70,012

Fonte: Elaboração dos Autores

Três fatores são extraídos carregando 70,012% da variância. A literatura recomenda 60% como mínimo de variância explicada. Todavia, essas cargas podem melhorar com exame detalhado das comunalidades. A variável “Roubo de Veículo” está abaixo do limite recomendado pela literatura (0,500), ou seja, essa é variável é pouco explicada pelos fatores extraídos. Hair et al (2005) recomenda a reestimação do modelo sem essas variáveis. A tabela apresenta as alterações na extração dos fatores com a retirada da variável com baixa comunalidade.

**Tabela 04 - Soma do Quadrado das Cargas Extraídas**

Componente	Total	Variância (%)	Variância Acumulada (%)
1	4,519	45,189	45,189
2	1,987	19,866	65,055
3	1,016	10,156	75,212

Fonte: Elaboração dos Autores

Como previsto, a variância carregada pelo primeiro fator, junto com a variância explicada pelo total de fatores aumentam (+ 7,427%). Todavia, o número de fatores extraídos continua o mesmo<sup>14</sup>. Em resumo, esses resultados evidenciam a multidimensionalidade da criminalidade dos estados brasileiros. Ou seja, dificilmente pode-se reduzir as diferentes modalidades de crime a um indicador.

### *Rotação dos Fatores*

Aqui a pretensão é comparar a rotação dos fatores a partir de um método ortogonal (Varimax) e um oblíquo (*Direct Oblimin*). A principal finalidade da rotação dos fatores é simplificar a solução oferecida pela análise fatorial. Ou seja, espera-se tornar mais clara a contribuição de cada variável na construção dos fatores. Precisamente, a estrutura torna-se mais simples na medida em que poucas ou nenhuma variável contribui com cargas acima de 0,40 para mais de um fator (HAIR ET AL, 2005). As tabelas abaixo apresentam os resultados.

**Tabela 05 - Rotação dos Fatores**

Variáveis	Varimax			<i>Direct Oblimin</i>		
	1	2	3	1	2	3
Estupro	0,369	<b>0,802</b>	0,136	<b>0,461</b>	,217	,866
Tentativa de Estupro	0,160	<b>0,872</b>	0,007	0,274	,067	,882
Tentativa de Homicídio	<b>0,529</b>	0,275	0,303	0,535	,377	,397
Homicídio	0,242	-0,319	<b>0,837</b>	0,131	,845	-,177
Lesão Corporal	0,380	<b>0,771</b>	0,038	<b>0,475</b>	,120	,827
Porte Ilegal de Armas	<b>0,805</b>	0,231	0,206	0,810	,312	,394
Roubo	-0,446	0,202	<b>0,517</b>	-0,454	,469	,172
Latrocínio	0,053	0,231	<b>0,856</b>	0,016	,866	,328
Tráfico de Entorpecentes	<b>0,819</b>	0,288	-0,067	0,852	,046	,423
Porte de Entorpecentes	<b>0,806</b>	0,349	-0,072	0,848	,043	,480

Fonte: Elaboração dos Autores

<sup>14</sup> Para não ficar apenas com o “Roubo a Instituição Financeira” decidimos utilizar a variável roubo – taxa de roubo de veículo e outros por 100 mil habitantes. Vale dizer a comunalidade dessa nova variável 0,508 os três fatores gerados nesse modelo carrega 72,05%.

**Tabela 06** - Matriz de Correlação de Fatores

Componentes	1	2	3
1	1,000	,047	,301
2	,047	1,000	,175
3	,301	,175	1,000

Fonte: Elaboração dos Autores

Pelo critério da simplificação, a rotação mais adequada é a Varimax (Ortogonal). Nela, nenhuma variável contribui positivamente acima do patamar (0,40) para mais de um fator. Contrariamente, na rotação por *direct oblimin* duas variáveis (“Estupro” e “Lesão Corporal”) apresentam estruturas complexas. Adicionalmente, a correlação entre os fatores é fraca<sup>15</sup>.

Em resumo, os resultados apontam a abordagem *varimax* como mais adequada. Metodologicamente, isso significa que os fatores são independentes entre si. Precisamente, o componente 1 é composto por “Tentativa de Homicídio”, “Tráfico de Entorpecentes”, “Porte de Entorpecentes” e “Porte Ilegal de Armas”. Já o componente 2 é composto por “Estupro”, “Tentativa de Estupro” e “Lesão Corporal”. Por fim, o componente 3 reúne as variáveis “Homicídio”, “Latrocínio” e “Roubo”<sup>16</sup>.

### *Redução dos Dados*

Normalmente, emprega-se a análise fatorial de componentes principais para escolher variáveis substitutas de um conjunto de outras. Para tanto, pode-se: 1) escolher a variável que mais contribui para um fator ou 2) trabalhar com os escores dos fatores extraídos. De acordo com Hair et al (2005), o pesquisador deve verificar a correspondência entre as duas estratégias. Como é muito comum a construção de *Rankings*, optamos por utilizar as duas estratégias para separar os estados brasileiros por níveis de criminalidade e em seguida comparar as classificações<sup>17</sup>. As tabelas abaixo apresentam a resposta.

---

<sup>15</sup> As correlações com magnitude  $\leq 0,40$  são consideradas fracas (ver DANCEY e REIDY, 2006).

<sup>16</sup> As variáveis estão listadas pela ordem de contribuição para seu principal fator.

<sup>17</sup> Todas as classificações seguiram o seguinte procedimento: 1) os casos foram agrupados via análise de cluster, abordagem *Two Step Clusters*, em três grupos e por distância Euclidiana; 2) primeiro, utilizamos o componente como parâmetro e distinguimos as categorias (alto, médio e baixo) a partir das médias dos

**Tabela 07 – Classificação dos Estados por Nível de Criminalidade (1ª Dimensão)<sup>a</sup>**

Grupos		Tráfico de Entorpecentes					
		Alto	%	Médio	%	Baixo	%
Componente 1	Alto	10	<b>66,67</b>	5	33,35	0	0,00
	Médio	2	7,69	19	<b>73,07</b>	5	19,2
	Baixo	0	0,00	2	12,50	14	<b>87,50</b>

Fonte:Elaboração dos Autores  $\chi^2 = 49,673$  p = 0,000 <sup>a</sup> Nível de Confluência = 75,44 %

**Tabela 08 – Classificação dos Estados por Nível de Criminalidade (2ª Dimensão)<sup>a</sup>**

Grupos		Tentativa de Estupro					
		Alto	%	Médio	%	Baixo	%
Componente 2	Alto	4	<b>100,00</b>	0	0,00	0	0,00
	Médio	0	0,00	10	<b>37,03</b>	17	62,97
	Baixo	0	0,00	0	0,00	26	<b>100,00</b>

Fonte:Elaboração dos Autores  $\chi^2 = 69,765$  p = 0,000 <sup>a</sup> Nível de Confluência = 70,02 %

**Tabela 09 – Classificação dos Estados por Nível de Criminalidade (3ª Dimensão)<sup>a</sup>**

Grupos		Latrocínio					
		Alto	%	Médio	%	Baixo	%
Componente 3	Alto	1	<b>50,00</b>	1	50,00	0	0,00
	Médio	0	0,00	16	<b>59,23</b>	11	40,74
	Baixo	0	0,00	1	3,57	27	<b>96,43</b>

Fonte:Elaboração dos Autores  $\chi^2 = 48,997$  p = 0,000 <sup>a</sup> Nível de Confluência = 77,19 %

Os estados foram classificados nas três dimensões de criminalidade a partir dos escores dos fatores e a partir das variáveis que mais contribuíram para criação deles<sup>18</sup>. A média de confluência das classificações é 74,217%<sup>19</sup>. Mais que isso, existe uma dependência estatisticamente significativa entre as classificações, seja qual for a dimensão analisada (p = 0,000). Em conjunto isso significa que as duas estratégias apresentam resultados aproximados. Porém, a análise em separado das tabelas revela em

---

grupos; 3) procedimento semelhante foi realizado tendo a variável que mais contribui para o componente como parâmetro. Por questão de espaço não informamos detalhes sobre as classificações, mas disponibilizamos a base de dados para que seja possível replicar os resultados, ver em: <http://sigobr.com/>  
<sup>18</sup> A série inclui os 27 Estados brasileiro em 3 anos de observação totalizando 81 unidades de análise. Todavia, devido a ausência de dados só foi possível classificar 57 unidades. Vale informar as etapas de sumarização e rotação dos fatores também foram realizadas imputando dados ausentes a partir do método selecionado pelo pacote estatístico. Os resultados são semelhantes a única diferença que a variável Roubo precisou ser substituída pela variável Roubo a Instituições Financeiras para manter a estrutura de sumarização dos dados (três componentes). A análise com dados imputados pode ser consultada nos anexos.

<sup>19</sup> O nível de confluência corresponde ao percentual de casos previsto de modo semelhante pelas duas classificações em relação ao total de casos.



que dimensão isso é mais verdadeiro. Sem dúvida a substituição é menos recomendada na dimensão 2 (componente 2 pela “Tentativa de Estupro”). Nela a divergência entre as classificações é maior e mais concentrada.

Por fim, vale comparar a classificação dos estados por composição dos fatores com variáveis arbitrariamente escolhidas. No Brasil, é comum dentro e fora da Academia a utilização da ocorrência de homicídios como medida de criminalidade. Mais que isso, é comum a criação de *rankings* dos estados baseado nessa variável. Dessa forma, cabe avaliar qual a convergência entre esses ordenamentos. As tabelas abaixo apresentam os resultados.

**Tabela 10** – Classificação dos Estados por Nível de Criminalidade (1ª Dimensão)<sup>a</sup>

Grupos		Homicídio					
		Alto	%	Médio	%	Baixo	%
Componente 1	Alto	2	<b>13,33</b>	8	53,33	5	33,33
	Médio	9	34,62	11	<b>42,08</b>	6	23,08
	Baixo	2	12,50	5	31,25	9	<b>56,25</b>

Fonte:Elaboração dos Autores  $\chi^2 = 6,947$   $p = 0,139$  <sup>a</sup> Nível de Confluência = 38,59%

**Tabela 11** – Classificação dos Estados por Nível de Criminalidade (2ª Dimensão)<sup>a</sup>

Grupos		Homicídio					
		Alto	%	Médio	%	Baixo	%
Componente 2	Alto	1	<b>25,00</b>	1	25,00	2	50,00
	Médio	1	3,70	16	<b>59,25</b>	10	37,03
	Baixo	11	42,30	7	26,92	8	<b>30,76</b>

Fonte: Elaboração dos Autores  $\chi^2 = 12,673$   $p = 0,013$  <sup>a</sup> Nível de Confluência = 43,86 %

**Tabela 12** – Classificação dos Estados por Nível de Criminalidade (3ª Dimensão)<sup>a</sup>

Grupos		Homicídio					
		Alto	%	Médio	%	Baixo	%
Componente 3	Alto	2	<b>100,00</b>	0	0,00	0	0,00
	Médio	10	37,04	15	<b>55,56</b>	2	7,41
	Baixo	1	3,57	9	32,14	18	<b>64,28</b>

Fonte: Elaboração dos Autores  $\chi^2 = 28,228$   $p = 0,000$  <sup>a</sup> Nível de Confluência = 61,403 %

O nível de confluência médio entre as classificações é de apenas 47,95%. Além disso, a comparação com o componente 1 aponta para independência das classificações ( $p > 0,05$ ). Em conjunto, esses resultados sinalizam que os *rankings* organizados pela

ocorrência de homicídio pode não ser um bom indicador do registro de outras modalidades de crime. Em outros termos, apontam que não há um compartilhamento de variância entre homicídio e as demais modalidades de crimes a ponto de se tomar um pelos demais.

### *Validação da Medida*

Para avaliar a validade e a confiabilidade do indicador de criminalidade utilizaremos três procedimentos: 1) análise da matriz de correlação residual; 2) fracionamento da amostra 3) validade e confiabilidade da medida e 4) perspectiva confirmatória.

### *Análise da matriz de correlação residual*

A matriz de correlação residual é uma medida do ajuste do modelo de rotação. Precisamente, ela corresponde à diferença entre a matriz de correlação original e a matriz de correlação reproduzida<sup>20</sup>. Quanto menor a diferença, menor a magnitude dos resíduos e mais bem ajustado é o modelo. De acordo com a literatura, o total de correlações residuais acima de 0,05 não deve ultrapassar 50% (BROWN, 2006; TABACHNICK e FIDELL, 2012). O quadro abaixo representa o resultado da matriz análise de correlação residual do indicador de criminalidade.

**Quadro 03 – Matriz de Correlação Residual**

	R1	R2	R3	R4	R5	R6	R7	R8	R9
R2	-0,099*								
R3	-0,012	<b>-0,113*</b>							
R4	-0,078*	<b>-0,101*</b>	-0,001						
R5	-0,042	0,066*	-0,067*	-0,024					
R6	-0,018	0,002	0,045	0,018	<b>-0,102*</b>				
R7	<b>-0,177*</b>	-0,040	-0,003	-0,010	-0,037	-0,069*			
R8	0,052*	0,000	-0,035	-0,063*	0,021	0,014	0,047		
R9	-0,063*	-0,017	0,046	0,046	-0,042	-0,021	-0,051*	-0,073*	
R10	-0,200*	-0,060*	0,060*	<b>0,107*</b>	-0,016	0,041	0,080*	<b>-0,170*</b>	<b>-0,165*</b>

R1 (Tentativa de Homicídio); R2 (Porte Ilegal de Armas); R3 (Tráfico de Entorpecentes); R4 (Porte de Entorpecentes); R5 (Estupro); R6 (Tentativa de Estupro); R7 (Lesão Corporal); R8 (Homicídio); R9 (Latrocínio); R10 (Roubo) \*Correlações com valores absolutos acima de 0,05, ao todo são 21 (46,67%).

<sup>20</sup> Baseada na correlação entre as variáveis originais e as cargas extraídas de cada uma delas.

No quadro acima é preciso analisar o sinal e a magnitude dos coeficientes. O sinal negativo significa que a correlação original entre as variáveis é menor que a correlação reproduzida após a extração dos fatores. Essa situação ocorre na maioria dos casos (30; 66,67%). Todavia, o mais relevante é a magnitude dessas diferenças. Em 21 (46,67%) casos a magnitude está acima do recomendado (0,05). Chama a atenção os coeficientes acima de 0,1 (07; 15,56%), a maioria deles envolvem os indicadores “Roubo” (03) e Porte Ilegal de Armas (02). Metodologicamente, esse resultado significa que é preciso observar com reservas a contribuição dessas variáveis para construção dos fatores.

#### *Fracionamento da Amostra*

Outro procedimento indicado para validar os resultados é replicar a análise. Aqui decidimos fracionar a amostra a partir do nível de cobertura dos dados nos estados<sup>21</sup>. Precisamente, selecionamos os estados com cobertura  $\geq 75,00\%$ <sup>22</sup>. O objetivo é observar a influência dessa variação e verificar se as etapas de extração e rotação dos fatores sofrem alterações. As tabelas abaixo ajudam na resposta.

**Tabela 13 - Comunalidades<sup>23</sup>**

Variáveis	Componentes Principais	
	Inicial	Extração
Estupro	1,000	0,795
Tentativa de Estupro	1,000	0,783
Tentativa de Homicídio	1,000	<b>0,438</b>
Homicídio	1,000	0,861
Lesão Corporal	1,000	0,701
Porte Ilegal de Armas	1,000	0,741
Roubo	1,000	<b>0,439</b>
Latrocínio	1,000	0,840
Tráfico de Entorpecentes	1,000	0,751
Porte de Entorpecentes	1,000	0,777

Fonte: Elaboração dos Autores

**Tabela 14 - Soma do Quadrado das Cargas Extraídas**

<sup>21</sup> Definir e referenciar esse nível.

<sup>22</sup> De acordo com o Fórum Brasileiro de Segurança Pública a cobertura da coleta de dados varia entre os estados. Alguns cobrem todo território enquanto outros cobrem menos de 50% dos seus municípios.

<sup>23</sup> Adequação da Amostra - KMO = 0,713 e BTS = 271,857 df = 45 e p = 0,000

Componentes Principais			
Componentes	Total	Variância (%)	Variância Acumulada (%)
1	3,970	39,700	39,700
2	1,793	17,926	57,625
3	1,363	13,634	71,260

Fonte: Elaboração dos Autores

**Tabela 15** - Rotação dos Fatores

Variáveis	Varimax		
	1	2	3
Estupro	<b>0,841</b>	0,279	0,098
Tentativa de Estupro	<b>0,878</b>	0,064	-0,088
Tentativa de Homicídio	0,414	0,420	0,301
Homicídio	-0,360	0,105	<b>0,849</b>
Lesão Corporal	<b>0,773</b>	0,321	-0,013
Porte Ilegal de Armas	0,250	<b>0,784</b>	0,255
Roubo	0,161	-0,544	<b>0,341</b>
Latrocínio	0,220	-0,113	<b>0,883</b>
Tráfico de Entorpecentes	0,299	<b>0,810</b>	-0,068
Porte de Entorpecentes	0,373	<b>0,797</b>	-0,050

Fonte: Elaboração dos Autores

Quanto às comunalidades, os indicadores “Tentativa de Homicídio” e “Roubo” têm patamares abaixo do recomendado pela literatura (0,50) (HAIR ET AL, 2005)<sup>24</sup>. Entretanto, o total de variância explicada é 71,260% (apenas 1,782% a mais que o modelo original)<sup>25</sup>. O número de componentes extraídos é o mesmo. Além disso, a composição deles é bastante parecida, é preciso apenas retirar a variável “Tentativa de Homicídio” do componente 1 e ignorar a contribuição negativa da variável “Roubo” ao componente 2. Em resumo, os resultados da análise com a amostra fracionada apresentam pouca diferença em relação à análise com a amostra completa.

#### *Validade e Confiabilidade*

Quanto maior é a validade, mas segura é a generalização dos resultados extraídos para toda a população. Aqui destacamos a validade convergente. Ou seja, a proporção de variância de cada item explicada pelo construto. Aqui tomaremos como

<sup>24</sup> A retirada das variáveis “Tentativa de Homicídio” e Roubo eleva para 81,953% a variância explicada, mas mantém o número e a estrutura de composição dos fatores.

<sup>25</sup> Comparação feita com o modelo reportado na tabela 04.

referência as cargas fatoriais da perspectiva exploratória<sup>26</sup>. A tabela abaixo apresenta a validade convergente do nosso índice:

**Tabela 16 – Validade Convergente**

<b>Componentes</b>	<b>Validade</b>
Componente 1	$\frac{2,248}{4} = \mathbf{0,562}$
Componente 2	$\frac{1,998}{3} = \mathbf{0,666}$
Componente 3	$\frac{1,700}{3} = \mathbf{0,567}$

Fonte: Elaboração dos Autores

Todos os componentes têm validade convergente acima do mínimo adequado (0,50) (HAIR ET AL, 2009). Isso significa que a variância explicada pelos componentes ultrapassa 50%. Isso é mais verdadeiro para o componente 2 (Estupro, Tentativa de Estupro e Lesão Corporal). Outro parâmetro de avaliação da medida é a sua confiabilidade. A tabela abaixo apresenta os dados para os três componentes.

**Tabela 17 – Confiabilidade**

<b>Construto</b>	<b>Confiabilidade</b>
Componente 1	$\frac{(2,248)^2}{(2,248)^2 + (1,752)^2} = \mathbf{0,622}$
Componente 2	$\frac{(1,998)^2}{(1,998)^2 + (1,000)^2} = \mathbf{0,799}$
Componente 3	

<sup>26</sup> Ver Tabela 05.

	$\frac{(1,700)^2}{(1,700)^2 + (1,299)^2} = \mathbf{0,631}$
--	--

Fonte: Elaboração dos Autores

Apenas o componente 2 apresenta uma confiabilidade acima do recomendado (0,70) (HAIR ET AL, 2009). Os demais apresentam coeficientes que inspira cuidado. Porém, de acordo com Hair et al (2009) confiabilidades entre 0,60 e 0,70 podem ser consideradas adequadas desde que o construto (componente) não apresente problemas de validade convergente. Como visto, esse é o caso dos componentes 1 e 3. Em resumo, a medida analisada apresenta níveis de validade e confiabilidade aceitáveis.

### *Perspectiva Confirmatória*

Segundo Brown (2006), a análise fatorial confirmatória é um tipo de modelo de equação estrutural cujo objetivo é detalhar a relação entre indicadores e variáveis latentes (fatores). Sua principal característica é ser orientada por uma hipótese. Aqui o objetivo é verificar o ajuste do modelo e conseqüentemente o nível de generalização dos resultados obtidos na construção do indicador de criminalidade. As tabelas abaixo apresentam o resultado.

**Tabela 18 - Comunalidades<sup>27</sup>**

Variáveis	Componentes Principais	
	Inicial	Extração
Estupro	,725	,892
Tentativa de Estupro	,624	,732
Tentativa de Homicídio	,460	,568
Homicídio	,573	,999
Lesão Corporal	,635	,713
Porte Ilegal de Armas	,660	,732
Roubo	,289	<b>,476</b>
Latrocínio	,521	,598
Tráfico de Entorpecentes	,775	,999
Porte de Entorpecentes	,765	,779

Fonte: Elaboração dos Autores

<sup>27</sup> Adequação da Amostra - KMO = 0,713 e BTS = 271,857 df = 45 e p = 0,000

**Tabela 19 - Soma do Quadrado das Cargas Extraídas**

Componentes Principais			
Componentes	Total	Variância (%)	Variância Acumulada (%)
1	2,753	27,532	27,532
2	1,682	16,824	44,356
3	1,928	19,278	63,634

Fonte: Elaboração dos Autores  $\chi^2 = 27,027$  Gl = 18 e p = 0,078

**Tabela 20 - Rotação dos Fatores**

Variáveis	Varimax		
	1	2	3
Estupro	<b>0,895</b>	0,260	0,086
Tentativa de Estupro	<b>0,770</b>	0,193	-0,082
Tentativa de Homicídio	0,389	0,348	0,301
Homicídio	-0,147	0,045	<b>0,988</b>
Lesão Corporal	<b>0,679</b>	0,402	0,045
Porte Ilegal de Armas	0,409	<b>0,570</b>	0,316
Roubo	-0,004	-0,006	<b>0,199</b>
Latrocínio	0,286	-0,001	<b>0,614</b>
Tráfico de Entorpecentes	0,189	<b>0,981</b>	-0,012
Porte de Entorpecentes	0,332	<b>0,802</b>	-0,006

Fonte: Elaboração dos Autores

Optamos por utilizar o método de *Generalized Last Square*<sup>28</sup>. No quesito comunalidade a abordagem confirmatória aponta para retirada da variável “Tentativa de Homicídio” do modelo (coeficiente abaixo do recomendado 0,500). No total, os três fatores extraídos explicam 63,634% da variância. O ajuste do modelo ( $\chi^2 = 27,027$  e p = 0,078) fica um pouco abaixo do recomendado ( $\chi^2 = 28,869$  p = 0,05). Por fim, os fatores extraídos apresentam a seguinte composição: fator 1 - “Estupro”, “Tentativa de Estupro” e “Lesão Corporal”; fator 2 - “Porte Ilegal de Armas”, “Tráfico de Entorpecentes” e “Porte de Entorpecentes” e o fator 3 - “Homicídio”; “Roubo”; “Latrocínio”. A variável “Tentativa de Homicídio” aparece com carga truncada (contribui quase igualmente para os três fatores).

<sup>28</sup> De acordo com Chen (2010), a *Generalized Last Square* flexibiliza pressupostos do modelo linear, precisamente: homocedasticidade, independência e normalidade dos resíduos. Como utilizamos amostra considera pequena e com possível dependência temporal das observações, optamos por esse método.

## CONCLUSÃO

A literatura especializada indica a bidimensionalidade da criminalidade. No Brasil, geralmente imprensa e academia tratam esses fenômenos como unidimensional. Mais precisamente, toma-se a ocorrência de homicídio como único indicador da criminalidade brasileira. Os resultados desse trabalho evidenciam a fragilidade dessa estratégia de mensuração. Primeiro, a análise indica que dificilmente pode-se reduzir as dez modalidades de crimes observadas a menos de três fatores. Na melhor das hipóteses, um fator explica 40,328% da variância total<sup>29</sup>. Segundo, Latrocínio explica mais variância do seu componente que homicídio (variável mais popular)<sup>30</sup>. Terceiro, o nível de confluência médio entre os *rankings* formados pelos fatores e pela ocorrência de homicídio é de apenas 47,95%. Ou seja, as classificações apresentam grandes divergências. Em outros termos, os dados apontam que não há um compartilhamento de variância entre homicídio e as demais modalidades de crimes a ponto de se tomar um pelos demais. Em conjunto, a validação da medida confirma os resultados da etapa exploratória. Isso significa que criminalidade brasileira oficialmente registrada no período apresenta três dimensões distintas.

Teoricamente, era possível prever a emergência de dois construtos. Eles são comumente chamados de crimes contra pessoa e crimes contra o patrimônio. Parece razoável chamar o componente 2 de crimes contra pessoa (composto por “Estupro”, “Tentativa de Estupro” e “Lesão Corporal”). Todavia, os componentes 01 e 03 fundem o que originalmente seria crimes contra pessoa e patrimônio. Nesse quesito, o que mais surpreende é a formação de um componente reunindo “Homicídio”, “Latrocínio” e “Roubo”. O que sugere um componente econômico a ocorrência de homicídio no Brasil.

Nosso estudo tem um caráter exploratório. Não testamos formalmente nenhuma hipótese. Portanto, não fornecemos bases seguras para afirmações definitivas. Para chegar lá, será preciso no mínimo replicar esse estudo com outras bases de dados e ampliar a série temporal. Por enquanto, nossa esperança é apenas animar novos trabalhos nessa direção e motivar o acúmulo de conhecimento na área.

## ANEXOS

---

<sup>29</sup> Ver tabela 03.

<sup>30</sup> Ver tabela 05 – Método Varimax .



Sumarização e rotação dos fatores com imputação de dados:

**Tabela 21** - Comunalidades<sup>31</sup>

Variáveis	Componentes Principais	
	Inicial	Extração
Estupro	1,000	0,799
Tentativa de Estupro	1,000	0,851
Tentativa de Homicídio	1,000	0,505
Homicídio	1,000	0,891
Lesão Corporal	1,000	0,836
Porte Ilegal de Armas	1,000	0,734
Roubo	1,000	0,659
Latrocínio	1,000	0,712
Tráfico de Entorpecentes	1,000	0,781
Porte de Entorpecentes	1,000	0,836

Fonte: Elaboração dos Autores

**Tabela 22** - Soma do Quadrado das Cargas Extraídas

Componentes Principais			
Componentes	Total	Variância (%)	Variância Acumulada (%)
1	4,142	41,416	41,416
2	2,122	21,217	62,633
3	1,294	12,940	75,573

Fonte: Elaboração dos Autores  $\chi^2 = 27,027$  Gl = 18 e p = 0,078

**Tabela 23** - Rotação dos Fatores

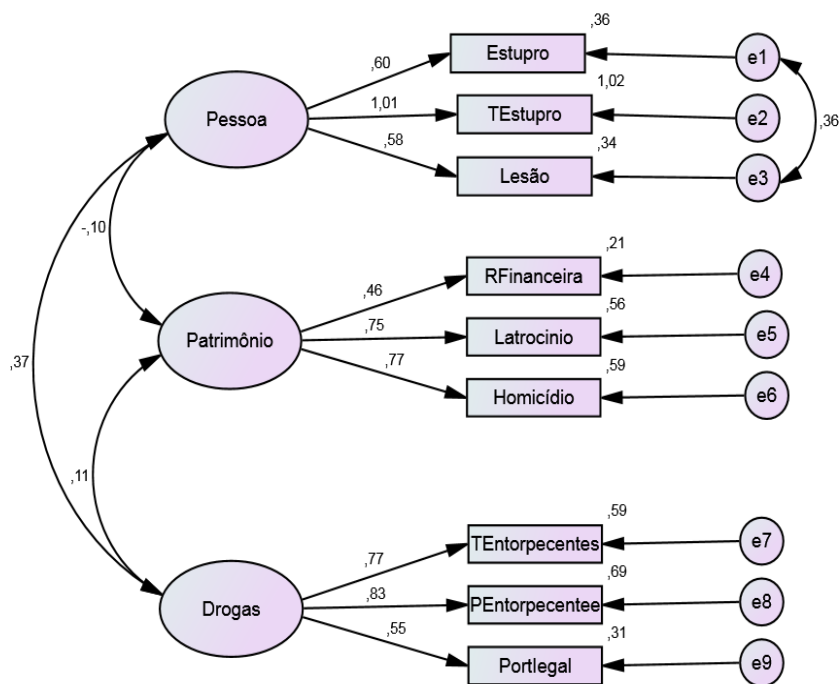
Variáveis	Varimax		
	1	2	3
Estupro	<b>0,800</b>	0,387	0,097
Tentativa de Estupro	<b>0,913</b>	0,117	0,065
Tentativa de Homicídio	0,379	<b>0,542</b>	0,260

<sup>31</sup> Adequação da Amostra - KMO = 0,713 e BTS = 271,857 df = 45 e p = 0,000

Homicídio	-0,284	0,093	<b>0,854</b>
Lesão Corporal	<b>0,842</b>	0,353	-0,54
Porte Ilegal de Armas	0,075	<b>0,764</b>	0,381
Roubo	0,321	-0,219	<b>0,712</b>
Latrocínio	0,109	0,163	<b>0,820</b>
Tráfico de Entorpecentes	0,230	<b>0,837</b>	-0,164
Porte de Entorpecentes	0,355	<b>0,850</b>	-0,120

Fonte: Elaboração dos Autores

Análise Fatorial Confirmatória



## REFERÊNCIAS BIBLIOGRÁFICAS

- ASHER, H. B. (1993), *Causal Modeling*. 2 ed. London: Sage.
- BEATO, F. C.; PEIXOTO, B. T. e ANDRADE, M. V. (2004), “Crime, Oportunidade e Vitimização”. *Revista Brasileira de Ciências Sociais*, vol. 19 n. 55.
- BECKER, G. S. (1968), “Crime and Punishment: an economic approach”. *Journal of Political Economy*. v. 76, n. 2.
- BLALOCK, H. M. (1979). The Presidential Address: Measurement and Conceptualization Problems: The Major Obstacle to Integrating Theory and Research. *American Sociological Review*, Vol. 44, No. 6 (Dec., 1979), pp. 881-894
- BLALOCK, H. M. Jr. (1967), “Causal inference, closed populations, and measures of association”. *American Political Science Review*, v. 61, p. 130-136.
- BLALOCK, H. M. Jr. (1971), *Causal Models in the Social Sciences*. Chicago: Aldine-Atherton.
- BROWN, T. A. (2006), *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.

- CONSTER, H. L.(1969), "Theory, Deduction, and Rules of Correspondence". *American Journal of Sociology*, n. 75 p. 245-263.
- COOMBS, C. (1964), *A Theory of Data*. New York: Wiley.
- DANCEY, Chirstine; REIDY, John (2006), *Estatística Sem Matemática para Psicologia: Usando SPSS para Windows*. Porto Alegre: Artmed.
- EHRlich (1973), "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation". *Journal of Political Economy*, v. 81, n. 3.
- EIDE, E. (1990), "Economics of Criminal Behavior". *Journal of Economic Law*.
- ENTORF, H. e SPENGLER, H. (1998), "Socio-economic and demographic factors of crime in Germany: evidence from panel data of the German states", ZEW Discussion Papers, n. 98-16
- GORDON, R. A. (1968), "Issue in Multiple Regression". *American Journal of Sociology*. v. 68 p. 195-204.
- GORSUCH, R. L. (1983), *Factor Analysis*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- GOULD, E. D.; WEINBERG, B. A. e MUSTARD, D. B. (2002), "Crime rates and local labor market opportunities in the United States: 1979-1997". *The Review of Economics and Statistics*. v. 84. n.1 p. 45-61.
- GUTTMAN, L. (1955), "The determinancy of factor score matrices with implications for five other basic problems of common factor theory". *British Journal of Statistical Psychology*, v.8, n. 65.
- HAIR, Joseph. F. ET AL (2005), *Análise Multivariada de Dados*. 6. Ed. Porto Alegre: Bookman.
- HAUSER, R. M. e GOLDBERGER, A. S. (1971), "The treatment of Unobservable Variable in Path Analysis". Costner, H. L. (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- HEERMAN (1963) E. F. (1963) "Univocal or orthogonal estimators of orthogonal factors". *Psychometrika*. v. 28, n.2 p.161-170.
- HEINEKE, J.M. (ed) (1978), *Economic Models of Criminal Behaviour*, Amsterdam, North-Holland, p. 391-402.
- HEISE, D. R. (1969), "Problems in Path Analysis and Causal Inference". In Borgatta, E. (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- JACOBSON, A. L. e LALU, N. M. (1974), "An Empirical and Algebraic Analysis of Alternative Techniques for Measuring Unobserved Variables". In\_\_\_ Blalock, H. M. (ed.) *Measurement in the Social Science. Theories and Strategies*, Chicago: AVC.
- KOLM, S.-C. (1973), "A Note on Optimal Tax Evasion". *Journal of Public Economics*, n. 2, p.

265-270.

LAZARSELD, P. F. e ROSENBERG, M. (Eds) (1955), Introduction to The Language of Social Research. New York: Free Press.

PETER, P. e TAYLOR, J. B. (1977), "The Deterrent Effect of Capital Punishment: Another View". *The American Economic Review*, v.67, n.3, pp. 445-451.

ROBERTS, F. S. (2009), *Measurement Theory: with application to decisionmaking, utility, and Social Science*. Cambridge: Cambridge University Press.

SINGH, B. (1973), 'Making Honesty the Best Policy', *Journal of Public Economics*, p257-263.

STIGLER, J. G. (1970), "The Optimum Enforcement of Laws". *Journal of Political Economy*. n. 78.

SULLIVAN, J. L. (1974), "Multiple indicators: some criteria of selection". In \_\_\_ Blalock, H. M. (ed.) *Measurement in the Social Science. Theories and Strategies*, Chicago: AVC.

TABACHNICK, B. e FIDELL, L. (2007), *Using Multivariate Statistics*, 5ed., Boston: Allyn e Bacon.

THOMPSON, Bruce (2004), *Exploratory and Confirmatory Factor Analysis: understanding concepts and applications*. Washington, DC: American Psychological Association.

WOLPIN, Kenneth I. (1978), "Capital Punishment and Homicide in England: A Summary of Results". *American Economic Review. Papers and Proceedings*, 422-427.

ZELLER, Richard A. e CARMINES, Edward G. (1980), *Measurement in the social sciences: the link between theory and data*. Cambridge, Cambridge University Press.