



Indicador Brasileiro de Criminalidade

José Alexandre da Silva Júnior¹ - <https://orcid.org/0000-0001-6600-7385>

Ranulfo Paranhos^{2*} - <https://orcid.org/0000-0001-7017-1540>

Willber Nascimento³ - <https://orcid.org/0000-0002-2257-8108>

Rodrigo Lins⁴ - <https://orcid.org/0000-0002-4521-3806>

Luís Roberto Nascimento de Oliveira⁵ - <https://orcid.org/0009-0006-3374-9722>

¹ Ministério do Desenvolvimento Social, Governo Federal, Brasília, Brasil

Email: jose.alexandre@mds.gov.br

² Instituto de Ciências Sociais, Univ. Fed. de Alagoas, Maceió, Brasil

Email: ranulfo.paranhos@ics.ufal.br

* corresponding author

³ Programa de Pós-Graduação em Sociologia, Univ. Fed. de Alagoas, Maceió, Brasil

Email: nascimentowillber@gmail.com

⁴ Instituto de Ciências Sociais, Univ. Fed. de Alagoas, Maceió, Brasil

Email: rodrigo.lins@ics.ufal.br

⁴ Instituto de Ciências Sociais, Univ. Fed. de Alagoas, Maceió, Brasil

Email: luís.oliveira@ics.ufal.br

Annals of Data Analysis

<https://periodicos.ufpe.br/revistas/ADA/>

Copyright © 2026 by author(s).

This work is licensed under the Creative Commons Attribution International License CC BY-NC-ND 4.0

<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>



Resumo

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.

Keywords

Mensuração; criminalidade; análise fatorial; homicídio

1. Introdução

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; 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; Lalu, 1974). Em outras palavras, ambas tendem a criar medidas com baixa validade e confiabilidade (Zeller; Carmines, 1980). Curiosamente, a literatura especializada nacional não procurou investigar o impacto dessa tendência¹. 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.

2. Estratégias de Mensuração

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. O quadro abaixo ajuda a compará-las:

Quadro 01 – Comparação das Estratégias de Mensuração

Estratégia	Fenômeno	Vantagem	Desvantagem
Indicador Único	Simples (univariado) e unidimensional.	Inteligível e mais simples de reproduzir.	Mais ausência de dados; Sujeito a erros não especificados e a baixa validade.

¹ ANTES DA REALIZAÇÃO DESSE TRABALHO FOI REALIZADA UM 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.

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 ambiguidades	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 seja, 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 está na simplicidade da medida. Todavia, corre-se o risco de ignorar a multidimensionalidade dos fenômenos e, conseqüentemente, elaborar uma medida contaminada por erros.

No caso da construção de um índice o intuito é mensurar fenômenos multivariados. 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; Lalu, 1974). De acordo com Hauser e Goldberger (1971), a estratégia mais comum é a mera adição dos itens sem atribuição de pesos. 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; 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 (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; Goldberger, 1971).

3. Mensurando Criminalidade

No Brasil, a maioria dos trabalhos utiliza a taxa de homicídios como indicador de criminalidade. Implícita ou explicitamente assume-se que a criminalidade é um fenômeno simples e unidimensional². 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). Essa técnica serve principalmente para o propósito de sintetizar padrões correlacionais existentes entre as variáveis observadas, possibilitando a condensação de conjuntos extensos de variáveis em um número reduzido de fatores (Tabachnick; Fidell, 2012). Portanto, a técnica tem três funções básicas: (1) sumarização, (2) redução e (3) validação dos dados.

² 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; REIDY, 2006).

Sumarização dos Dados

O foco aqui é a correlação entre as variáveis que se pretende sumarizar. 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, incluindo o erro e variância específica de cada variável (Tabachnick; Fidell, 2012)³. 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. 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; Fidell, 2012). Existem dois tipos de rotação: ortogonal e oblíquo. Na rotação ortogonal, os fatores são independentes entre si, no modelo 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; Fidell, 2012). A *varimax* tende a redistribuir proporcionalmente a variância entre os fatores, tornando-os igualmente importantes⁴. 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. Portanto, a técnica aposta na dependência entre as dimensões.

Redução dos Dados

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. 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 comunicações são altas e (2) a relação entre as variáveis e o fator é clara (Guttman, 1955; Heermann, 1963; Gorsuch, 1983). 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). Aqui utilizaremos um método que gera escores não correlacionados (Anderson-Rubin). Em resumo, pretendemos contrapor três formas de mensurar a criminalidade: (1) variável metodologicamente escolhida; (2) variável arbitrariamente escolhida; e (3) composição de fatores.

Validação da Medida

³ O MÉTODO CRIA TANTOS COMPONENTES QUANTAS VARIÁVEIS HOVER. 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 EXTRAÍ O MÁXIMO DE VARIÂNCIA NÃO CORRELACIONADA COM O PRIMEIRO COMPONENTE” (TABACHNICK; FIDELL, 2012 P.640).

⁴ 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; Fidell, 2012).

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; Carmines, 1979). A validade é a proporção em que a medida mensura aquilo a que ela se propõe. A ausência de indicadores que contribuem para construção de dois fatores é sinal da confiabilidade da medida (Sullivan, 1974; Gordon, 1968).

Para avaliar a confiabilidade e validade das medidas construída por análise fatorial, precisamos de: (1) análise da matriz de correlação residual, (2) fracionamento da amostra e (3) perspectiva confirmatória. A matriz de correlação residual é uma medida da utilização da correlação entre as variáveis originais para construção dos fatores. A perspectiva confirmatória é o método mais direto para validar os resultados de uma análise fatorial. Aqui utilizaremos essa perspectiva a partir do método de máxima verossimilhança. De acordo com esse método procura estimar parâmetros que maximizem a probabilidade de observar resultados iguais, caso os dados sejam novamente coletados na mesma população (Brown, 2006; Thompson, 2004)⁵. Para tanto, precisa respeitar alguns pressupostos: (1) amostras grandes (assintóticas); (2) indicadores (variáveis observadas) contínuos; e com (3) distribuição normal. Respeitado os pressupostos, a técnica é adequada para testar a confiabilidade dos fatores gerados.

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>).

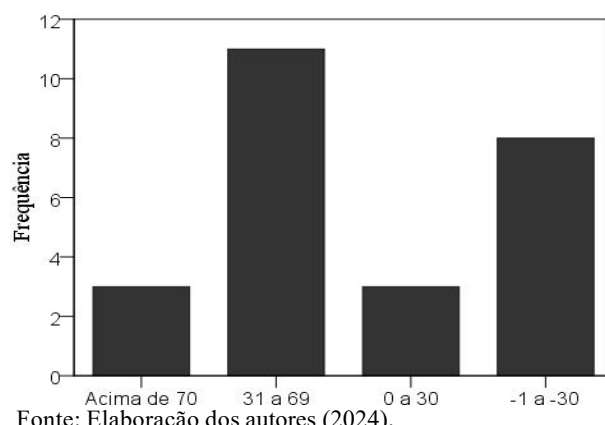
Fonte: Elaboração dos autores (2024).

4. Resultados

O primeiro passo da análise é o exame da matriz de correlação e de covariância das variáveis observadas. O gráfico abaixo mostra que a maior parte dos coeficientes tem magnitude acima de 0,30 (14; 56%), considerando apenas os coeficientes positivos com significância estatística ($p < 0,05$).

Gráfico 01 - Coeficientes de Correlação

⁵ 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).



Esse resultado é animador para a análise fatorial (Hair *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; Fidell, 2012). 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 Sphericity* (BTS) e do Kaiser-Meyer-Olkin (KMO). As tabelas abaixo apresentam os resultados.

Tabela 01 – Adequação da Amostra

Testes	Componentes Principais	
KMO	Maguinite	0,693
	χ^2 Aproximado	305,624
BTS	Grau de Liberdade	55
	p	0,000

Fonte: Elaboração dos autores (2024).

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	0,383
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 (2024).

Tabela 03 – Soma dos quadrados 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 (2024).

Três fatores são extraídos carregando 70,012% da variância. A literatura recomenda 60% como mínimo de variância explicada. 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 (2024).

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 mesmo⁶¹⁴. Em resumo, esses resultados evidenciam a multidimensionalidade da

⁶ 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%.

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

Fonte: Elaboração dos autores (2024).

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 (2024).

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 fraca⁷.

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”¹⁶.

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

⁷ As correlações com magnitude $\leq 0,40$ são consideradas fracas (ver Dancey; Reidy, 2006).

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 classificações¹⁷. As tabelas abaixo apresentam a resposta.

Tabela 07 – Classificação dos Estados por Nível de Criminalidade (1ª Dimensão)^a

Grupos		Tráfico de Entorpecentes					
		Alto	%	Médio	%	Baixo	%
Componente1	Alto	10	66,67	5	33,35	0	0,00
	Médio	2	7,69	19	73,07	5	19,2
	Baixo	0	0,00	2	12,50	14	87,50

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

Tabela 08 – Classificação dos Estados por Nível de Criminalidade (2ª Dimensão)^a

Grupos		Tentativa de Estupro					
		Alto	%	Médio	%	Baixo	%
Componente2	Alto	4	100,00	0	0,00	0	0,00
	Médio	0	0,00	10	37,03	17	62,97
	Baixo	0	0,00	0	0,00	26	100,00

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

Tabela 09 – Classificação dos Estados por Nível de Criminalidade (3ª Dimensão)^a

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

Fonte:Elaboração dos Autores $\chi^2 = 48,997$ p = 0,000 ^a 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 deles¹⁸. A média de confluência das classificações é 74,217%¹⁹. 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 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)^a

Grupos		Homicídio					
		Alto	%	Médio	%	Baixo	%
	Alto	2	13,33	8	53,33	5	33,33

Componente1	Médio	9	34,62	11	42,08	6	23,08
	Baixo	2	12,50	5	31,25	9	56,25

Fonte:Elaboração dos Autores $\chi^2=6,947$ $p=0,139$ ^aNível de Confluência = 38,59%

Tabela 11 – Classificação dos Estados por Nível de Criminalidade (2ª Dimensão)^a

Grupos		Homicídio					
		Alto	%	Médio	%	Baixo	%
Componente2	Alto	1	25,00	1	25,00	2	50,00
	Médio	1	3,70	16	59,25	10	37,03
	Baixo	11	42,30	7	26,92	8	30,76

Fonte: Elaboração dos Autores $\chi^2 = 12,673$ $p = 0,013$ ^aNível de Confluência = 43,86 %

Tabela 12 – Classificação dos Estados por Nível de Criminalidade (3ª Dimensão)^a

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

Fonte: Elaboração dos Autores $\chi^2 = 28,228$ $p = 0,000$ ^aNível 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 de outras modalidades de crime.

Análise da matriz de correlação residual

A matriz de correlação residual é uma medida do ajuste do modelo de rotação. Ela corresponde à diferença entre a matriz de correlação original e a matriz de correlação reproduzida²⁰. Quanto menor a diferença, menor a magnitude dos resíduos e mais bem ajustado é o modelo. O total de correlações residuais acima de 0,05 não deve ultrapassar 50% (Brown, 2006; Tabachnick; Fidell, 2012). O quadro abaixo representa o resultado da nossa matriz de correlação residual.

Quadro 03 – Matriz de Correlação Residual

	R1	R2	R3	R4	R5	R6	R7	R8	R9
R2	-0,099*								
R3	-0,012	-0,113*							
R4	-0,078*	-0,101*	-0,001						
R5	-0,042	0,066*	-0,067*	-0,024					
R6	-0,018	0,002	0,045	0,018	-0,102*				
R7	-0,177*	-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*	0,107*	-0,016	0,041	0,080*	-0,170*	-0,165*

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%).

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). 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. Precisamente, selecionamos os estados com cobertura $\geq 75,00\%$ ⁸. 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 14 - Soma do Quadrado das Cargas Extraídas

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 (2024).

Tabela 15 - Rotação dos Fatores

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

Fonte: Elaboração dos autores (2024).

⁸ 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.

Quanto às comunalidades, os indicadores “Tentativa de Homicídio” e “Roubo” têm patamares abaixo do recomendado pela literatura (0,50) (Hair *et al.*, 2005)²⁴. Entretanto, o total de variância explicada é 71,260% (apenas 1,782% a mais que o modelo original). 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.

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 referência as cargas fatoriais da perspectiva exploratória. A tabela abaixo apresenta a validade convergente do nosso índice:

Tabela 16 – Validade Convergente

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

Fonte: Elaboração dos autores (2024).

Todos os componentes têm validade convergente acima do mínimo adequado (0,50) (Hair *et al.*, 2005). 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

Construto	Confiabilidade
Componente 1	$\frac{(2,248)^2}{(2,248)^2 + (1,752)^2} = 0,622$
Componente 2	$\frac{(1,998)^2}{(1,998)^2 + (1,000)^2} = 0,799$
Componente 3	$\frac{(1,700)^2}{(1,700)^2 + (1,299)^2} = 0,631$

Fonte: Elaboração dos autores (2024).

Apenas o componente 2 apresenta uma confiabilidade acima do recomendado (0,70) (Hair *et al.*, 2005). 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 adequada 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 consequentemente o nível de generalização dos resultados obtidos na construção do indicador de criminalidade. As tabelas abaixo apresentam o resultado.

Tabela 18 - Comunalidade

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	,476
Latrocínio	,521	,598
Tráfico de Entorpecentes	,775	,999
Porte de Entorpecentes	,765	,779

Fonte: Elaboração dos autores (2024).

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

$\chi^2 = 27,027$ Gl = 18 e $p = 0,078$

Fonte: Elaboração dos autores (2024).

Tabela 20 - Rotação dos Fatores

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

Fonte: Elaboração dos autores (2024).

Optamos por utilizar o método de *Generalized Last Square*. 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 ($\chi^2 = 28,869$ e $p = 0,05$). Por fim, os fatores extraídos apresentam a seguinte composição: fator 1 - “Estupro”, “Tentativa de Estupro” e “Lesão Corporal”; fator 2 - “Porte

“Homicídio”, “Roubo”, “Latrocínio”. A variável “Tentativa de Homicídio” aparece com carga truncada (contribui quase igualmente para os três fatores).

5. Considerações Finais

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 total²⁹. Segundo, Latrocínio explica mais variância do seu componente que homicídio (variável mais popular)³⁰. 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.

Referências Bibliográficas

- Asher, H. B. (1993). *Causal modeling* (2nd ed.). Sage.
- Beato, F. C., Peixoto, B. T., & Andrade, M. V. (2004). Crime, oportunidade e vitimização. *Revista Brasileira de Ciências Sociais*, 19(55).
- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2).
- Blalock, H. M. (1979). The presidential address: Measurement and conceptualization problems: The major obstacle to integrating theory and research. *American Sociological Review*, 44(6), 881-894.
- Blalock, H. M., Jr. (1967). Causal inference, closed populations, and measures of association. *American Political Science Review*, 61, 130-136.
- Blalock, H. M., Jr. (1971). *Causal models in the social sciences*. Aldine-Atherton.

- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. The Guilford Press.
- Conster, H. L. (1969). Theory, deduction, and rules of correspondence. *American Journal of Sociology*, 75, 245-263.
- Coombs, C. (1964). *A theory of data*. Wiley.
- Dancey, C., & Reidy, J. (2006). *Estatística sem matemática para psicologia: Usando SPSS para Windows*. Artmed.
- Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy*, 81(3).
- Eide, E. (1990). Economics of criminal behavior. *Journal of Economic Law*.
- Entorf, H., & Spengler, H. (1998). Socio-economic and demographic factors of crime in Germany: Evidence from panel data of the German states. *ZEW Discussion Papers*, 98-16.
- Gordon, R. A. (1968). Issue in multiple regression. *American Journal of Sociology*, 68, 195-204.
- Gorsuch, R. L. (1983). *Factor analysis*. Lawrence Erlbaum Associates.
- Gould, E. D., Weinberg, B. A., & Mustard, D. B. (2002). Crime rates and local labor market opportunities in the United States: 1979-1997. *The Review of Economics and Statistics*, 84(1), 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*, 8, 65.
- Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (2005). *Análise multivariada de dados* (6th ed.). Bookman.
- Hauser, R. M., & Goldberger, A. S. (1971). The treatment of unobservable variable in path analysis. In H. L. Costner (Ed.), *Sociological methodology* (pp. 81-117). Jossey-Bass.
- Heerman, E. F. (1963). Univocal or orthogonal estimators of orthogonal factors. *Psychometrika*, 28(2), 161-170.
- Heineke, J. M. (Ed.). (1978). *Economic models of criminal behaviour* (pp. 391-402). North-Holland.
- Heise, D. R. (1969). Problems in path analysis and causal inference. In E. Borgatta (Ed.), *Sociological methodology* (pp. 38-73). Jossey-Bass.
- Jacobson, A. L., & Lalu, N. M. (1974). An empirical and algebraic analysis of alternative techniques for measuring unobserved variables. In H. M. Blalock (Ed.), *Measurement in the social science: Theories and strategies* (pp. 215-242). AVC.
- Kolm, S.-C. (1973). A note on optimal tax evasion. *Journal of Public Economics*, 2, 265-270.
- Lazarsfeld, P. F., & Rosenberg, M. (Eds.). (1955). *Introduction to the language of social research*. Free Press.
- Peter, P., & Taylor, J. B. (1977). The deterrent effect of capital punishment: Another view. *The American Economic Review*, 67(3), 445-451.
- Roberts, F. S. (2009). *Measurement theory: With application to decisionmaking, utility, and social science*. Cambridge University Press.
- Singh, B. (1973). Making honesty the best policy. *Journal of Public Economics*, 2, 257-263.

- Stigler, J. G. (1970). The optimum enforcement of laws. *Journal of Political Economy*, 78.
- Sullivan, J. L. (1974). Multiple indicators: Some criteria of selection. In H. M. Blalock (Ed.), *Measurement in the social science: Theories and strategies* (pp. 243-269). AVC.
- Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. American Psychological Association.
- Wolpin, K. I. (1978). Capital punishment and homicide in England: A summary of results. *American Economic Review: Papers and Proceedings*, 422-427.
- Zeller, R. A., & Carmines, E. G. (1980). *Measurement in the social sciences: The link between theory and data*. Cambridge University Press.

ANEXOS

Sumarização e rotação dos fatores com imputação de dados:

Tabela 21 - Comunalidades³¹

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 (2024).

Tabela 22 – Some 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	0,800	0,387	0,097
Tentativa de Estupro	0,913	0,117	0,065
Tentativa de Homicídio	0,379	0,542	0,260
Homicídio	-0,284	0,093	0,854
Lesão Corporal	0,842	0,353	-0,54
Porte Ilegal de Armas	0,075	0,764	0,381
Roubo	0,321	-0,219	0,712
Latrocínio	0,109	0,163	0,820
Tráfico de Entorpecentes	0,230	0,837	-0,164
Porte de Entorpecentes	0,355	0,850	-0,120

Fonte: Elaboração dos Autores

Análise Fatorial Confirmatória

