

Diretrizes para o Relato de Análise Fatorial Confirmatória em Psicologia Organizacional e do Trabalho

Guidelines for Reporting Confirmatory Factor Analysis in Organizational and Industrial Psychology

Directrices para la Redacción de Informes de Análisis Factorial Confirmatorio en Psicología Organizacional y del Trabajo

Opinião

Germano Gabriel Lima Esteves^{1,2}

<https://orcid.org/0000-0002-1851-4603>

E-mail: germanoesteves@unirv.edu.br

¹ Universidade de Rio Verde (UnirV), Rio Verde, Goiás (GO), Brasil

² Programa de Pós-graduação em Psicologia Social do Trabalho e das Organizações, Universidade Federal de Brasília (UnB), Brasília, Distrito Federal (DF), Brasil

Editoras Associadas Responsáveis:

Gardênia da Silva Abbad

<https://orcid.org/0000-0003-0807-3549>

Mary Sandra Carlotto

<https://orcid.org/0000-0003-2336-5224>

Laila Leite Carneiro

<https://orcid.org/0000-0001-7183-0501>

Como citar:

Esteves, G. G. L. (2026). Diretrizes para o Relato de Análise Fatorial Confirmatória em Psicologia Organizacional e do Trabalho. *Revista Psicologia: Organizações e Trabalho*, 26, e26935.

<https://doi.org/10.5935/rpot/e26935>



Resumo: A pesquisa em Psicologia Organizacional e do Trabalho (POT) utiliza instrumentos psicométricos para operacionalizar construtos latentes. Contudo, as pesquisas muitas vezes empregam medidas estabelecidas sem examinar se a estrutura de mensuração hipotetizada se mantém na amostra estudada. Essa prática pode comprometer a adequação das conclusões quando a relação entre construtos latentes e indicadores observados é assumida em vez de testada. Este artigo discute a importância da adequação dos modelos de mensuração às amostras, argumenta a favor do uso da Análise Fatorial Confirmatória (AFC) e fornece diretrizes para seu emprego. Para isso, delineamos os fundamentos conceituais que vinculam as evidências de validade baseada na estrutura interna e apresentamos recomendações para a apresentação da AFC em estudos empíricos. Além disso, enfatizamos o princípio da proporcionalidade metodológica, sugerindo que as expectativas quanto à avaliação da mensuração estejam alinhadas ao escopo de cada estudo. Assim, pretende-se contribuir para o aprimoramento do rigor metodológico e o desenvolvimento do conhecimento na pesquisa em POT.

Palavras-chave: análise fatorial confirmatória, psicometria, psicologia organizacional e do trabalho, validade, rigor metodológico.

Abstract: Research in Organizational and Work Psychology (OWP) uses psychometric instruments to operationalize latent constructs. However, studies often employ established measures without examining whether the hypothesized measurement structure holds in the sample studied. This practice may compromise the adequacy of conclusions when the relationship between latent constructs and observed indicators is assumed rather than tested. This article discusses the importance of the adequacy of measurement models in samples, argues in favor of the use of Confirmatory Factor Analysis (CFA), and provides guidelines for its use. To this end, we outline the conceptual foundations that link validity evidence based on internal structure and present recommendations for reporting CFA in empirical studies. In addition, we emphasize the principle of methodological proportionality, suggesting that expectations regarding the evaluation of measurement should be aligned with the scope of each study. Thus, the aim is to contribute to improving methodological rigor and advancing knowledge in OWP research.

Keywords: confirmatory factor analysis, psychometrics, organizational and work psychology, validity, methodological rigor.

Resumen: La investigación en Psicología Organizacional y del Trabajo (POT) utiliza instrumentos psicométricos para operacionalizar constructos latentes. Sin embargo, los estudios suelen emplear medidas establecidas sin examinar si la estructura de medición hipotética se mantiene en la muestra estudiada. Esta práctica puede comprometer la validez de las conclusiones cuando la relación entre los constructos latentes y los indicadores observados se da por sentada en lugar de someterse a prueba. Este artículo aborda la importancia de la adecuación de los modelos de medición a las muestras, aboga por el uso del Análisis Factorial Confirmatorio (AFC) y ofrece directrices para su aplicación. Para ello, esbozamos los fundamentos conceptuales que vinculan las evidencias de validez basadas en la estructura interna y presentamos recomendaciones para la aplicación del AFC en estudios empíricos. Además, hacemos hincapié en el principio de proporcionalidad metodológica, sugiriendo que las expectativas respecto a la evaluación de la medición estén alineadas con el alcance de cada estudio. De este modo, se pretende contribuir a la mejora del rigor metodológico y al desarrollo del conocimiento en la investigación sobre la POT.

Palabras clave: análisis factorial confirmatorio, psicometría, psicología organizacional y del trabajo, validez, rigor metodológico.

Introdução

Nas últimas décadas, a psicologia organizacional e do trabalho consolidou-se como um campo científico fortemente orientado por evidências empíricas derivadas de medidas psicológicas. Entretanto, a qualidade das conclusões produzidas depende não apenas do delineamento do estudo, mas também da adequação estrutural da medida utilizada. Mesmo quando o estudo não tem como objetivo a análise de parâmetros psicométricos, a consistência do modelo teórico de mensuração permanece uma condição necessária para sustentar as interpretações.

Embora ainda não exista uma exigência formal universal para que todos os estudos reportem evidências de validade e estimativas de confiabilidade das medidas utilizadas, periódicos como *Psychological Assessment*, *Journal of Personality Assessment* e *Assessment* enfatizam que pesquisas que desenvolvem, adaptam ou aplicam instrumentos em novas populações devem apresentar evidências adequadas de validade baseadas na estrutura interna, bem como estimativas de consistência interna e de outras propriedades psicométricas relevantes.

No campo da Psicologia Organizacional e do Trabalho (POT), a leitura de estudos publicados em periódicos como o *Journal of Vocational Behavior* e o *Journal of Occupational and Organizational Psychology* evidencia que a adequação do modelo teórico de mensuração aos dados empíricos constitui condição recorrente para a sustentação de inferências complexas. Tal aspecto torna-se particularmente saliente em pesquisas que envolvem mediação e moderação, nas quais a adequação estrutural dos construtos latentes constitui uma premissa fundamental para a interpretação dos resultados.

Considerando a centralidade das medidas psicológicas na pesquisa em POT, argumenta-se que estudos empíricos que utilizam instrumentos psicométricos deveriam reportar evidências de validade baseadas na estrutura interna estimada na própria amostra investigada, preferencialmente por meio da Análise Fatorial Confirmatória (AFC), ainda que o objetivo do estudo não seja a investigação de parâmetros psicométricos da medida. Assim, este artigo discute a importância de estimar evidências de validade baseadas na estrutura interna por meio da AFC em pesquisas de POT, mesmo quando o objetivo do estudo não é a obtenção de evidências de validade de instrumentos. Especificamente, o artigo: (a) discute o papel da AFC na sustentação das inferências produzidas pela pesquisa; (b) apresenta recomendações sobre quais informações devem ser reportadas na AFC; (c) apresenta recomendações para um relato transparente e proporcional ao escopo do estudo; e (d) apresenta implicações para a pesquisa em Psicologia Organizacional e do Trabalho.

O papel da AFC na sustentação das inferências produzidas pelas pesquisas

Para além de operacionalizar conceitos científicos, as medidas psicológicas (p. ex., inventários, escalas) devem ser coerentes com o contexto sociocultural e organizacional em que serão utilizadas. A formulação de construtos para explicar fenômenos organizacionais ou de trabalho em um país ou região pode não manter a adequação explicativa em contextos distintos. Conseqüentemente, instrumentos desenvolvidos para quantificar tais construtos podem apresentar propriedades psicométricas satisfatórias em uma população específica, mas não manter o desempenho em outros contextos culturais ou organizacionais (Prince, 2008).

Nesse sentido, a Psicologia, de modo mais específico a Psicometria, tem enfatizado a necessidade de obtenção de evidências de validade e estimativas de confiabilidade fundamentadas na população e no contexto em que a medida é aplicada, reconhecendo que tais propriedades não são inerentes ao instrumento em si, mas às interpretações produzidas a partir de seus escores (*American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014*).

Tradicionalmente, a validade foi concebida como uma propriedade relativamente estável da medida, frequentemente organizada em três categorias distintas: validade de conteúdo, validade de critério e validade de construto (Cronbach & Meehl, 1955). Contudo, esse entendimento foi progressivamente reformulado em direção a uma concepção unificada e cumulativa. Nessa perspectiva contemporânea, a validade passou a ser compreendida como um argumento fundamentado em múltiplas fontes de evidências que sustentam a interpretação dos escores, incluindo evidências baseadas na estrutura interna, nas relações com variáveis externas e no conteúdo do instrumento (AERA et al., 2014).

Esse modelo integrativo desloca o foco da medida em si para a adequação das interpretações produzidas em contextos específicos de aplicação, ressaltando também a necessidade de estimativas apropriadas de precisão, como a consistência interna (AERA et al., 2014). Diretrizes recentes da *American Psychological Association* (2020) reforçam essa compreensão ao enfatizar a responsabilidade do pesquisador quanto à fundamentação empírica das inferências realizadas.

Entre as evidências de validade associadas às medidas psicológicas, destacam-se aquelas baseadas na estrutura interna, usualmente investigadas por meio da Análise Fatorial Exploratória (AFE), da Análise Fatorial Confirmatória (AFC) e de modelos fundamentados na Teoria de Resposta ao Item (TRI) (AERA et al., 2014). A AFE compreende um conjunto de técnicas multivariadas que têm como propósito identificar a estrutura latente subjacente às variáveis observadas, buscando representar de forma parcimoniosa as inter-relações entre os itens (Brown, 2015; Hair et al., 2009).

De modo geral, é empregada quando o pesquisador pretende explorar como os itens de um instrumento se organizam na amostra analisada, seja no desenvolvimento de um novo inventário ou escala, seja em processos de adaptação cultural (Beaton et al., 2000; DeVellis, 2017; Fabrigar et al., 1999).

Contudo, mesmo em estudos nos quais medidas psicológicas são utilizadas para investigar fenômenos, e não para estimar parâmetros psicométricos, permanece fundamental assegurar que a estrutura teórica proposta para o instrumento encontre respaldo empírico na amostra estudada. É nesse contexto que a AFC pode se configurar como a estratégia mais adequada para obter evidências sobre a adequação dos dados à estrutura fatorial das medidas psicológicas. Apesar de guardar semelhanças com a AFE, a AFC se diferencia por depender da estrutura teórica que o pesquisador fornecer para a realização da análise. Na prática, isso significa que o pesquisador está “dizendo” à análise quais itens pertencem a quais fatores (Brown, 2015; Kline, 2016). Dessa forma, o nível de clareza sobre a adequação da utilização dos inventários e/ou escalas psicológicas naquela população e contexto deve se dar tanto pela natureza teórica do que se quer avaliar quanto pela clareza empírica.

No contexto em que problemas em modelos estatísticos frequentemente decorrem da inadequação entre o modelo teórico de mensuração e os dados empíricos observados (Brown, 2015), o ajuste adequado da estrutura fatorial por meio da AFC não apenas fornece evidências sobre a pertinência da medida à população e ao contexto investigados, mas também repercute diretamente na plausibilidade das inferências derivadas dos escores obtidos (AERA et al., 2014).

Recomendações sobre quais informações devem ser reportadas na AFC

O presente artigo discute a relevância de estimar e reportar evidências de validade baseadas na estrutura interna, por meio da Análise Fatorial Confirmatória (AFC), em pesquisas que utilizam instrumentos psicométricos na POT. As recomendações aqui apresentadas não devem ser interpretadas como um protocolo universal ou prescritivo para o relato de AFC. Pelo contrário, elas têm caráter orientador e visam estimular maior rigor psicométrico na área, preservando a autonomia dos pesquisadores para empregar abordagens metodológicas alternativas quando devidamente justificadas do ponto de vista teórico e empírico.

Considerando tais aspectos, é fundamental que pesquisadores que realizam AFC reportem de forma transparente as decisões analíticas adotadas, os procedimentos de estimação empregados e os índices globais de ajuste que permitam avaliar a adequação dos dados à estrutura interna teoricamente proposta para as medidas psicológicas utilizadas no estudo. Nesse sentido, com base nas recomendações de diferentes guias metodológicos sobre AFC (e.g.: Barrett, 2007; Bentler, 2007; Boomsma, 2000; Brown, 2015; Hooper et al., 2008; Jackson et al., 2009; MacCallum & Austin, 2000; McDonald & Ho, 2002), sugere-se que estudos que utilizam AFC reportem, de forma clara e objetiva, pelo menos as seguintes informações:

(i) **Método de estimação (*Estimation Method*)** – A escolha do método de estimação constitui uma decisão metodológica fundamentada nas características da medida e na natureza dos dados coletados. Em termos técnicos, o método de estimação corresponde ao procedimento estatístico utilizado para estimar os parâmetros do modelo (cargas fatoriais, variâncias, covariâncias e erros), de modo que a matriz de covariância implícita no modelo reproduza, com o menor grau possível de discrepância, a matriz observada nos dados (Brown, 2015; Jackson et al., 2009). O relato explícito do método de estimação é essencial, dado que diferentes procedimentos assumem pressupostos distintos (e.g., quanto à normalidade e ao nível de mensuração dos dados) e podem influenciar diretamente os valores dos parâmetros estimados, os índices de ajuste e o estatístico de qui-quadrado (χ^2), conforme evidenciado em estudos de simulação (Brown, 2015; Doğan, 2015; Fan et al., 1999; Powell & Schafer, 2001). A escolha do método de estimação deve estar alinhada com o nível de mensuração dos indicadores e com as propriedades de distribuição dos dados (Bollen, 1989; Brown, 2015; Kline, 2016). Fornecemos aqui uma breve descrição de alguns métodos de estimação que pode guiar a escolha do pesquisador (Ver Tabela 1). Contudo, descrições detalhadas dos métodos de estimação comumente usados na AFC podem ser encontradas em Brown (2015), Kline (2016) e Muthén e Muthén (2017).

Tabela 1

Métodos de estimação mais comuns na AFC

Método de Estimação	Tipo de dado recomendado	Características
<i>Maximum Likelihood (ML)</i>	Variáveis contínuas aproximadamente normais	Estimador mais utilizado; estimativas eficientes; permite comparação entre modelos via teste de razão de verossimilhança; bom desempenho em amostras moderadas ou grandes
<i>Robust Maximum Likelihood (MLR/MLM)</i>	Variáveis contínuas com violação de normalidade	Corrige erros padrão e estatística χ^2 ; mais robusto à assimetria e curtose; amplamente disponível em softwares SEM

Método de Estimação	Tipo de dado recomendado	Características
<i>Weighted Least Squares (WLS)</i>	Variáveis ordinais (e.g.: escalas Likert) com violação de normalidade	Utiliza correlações policóricas; adequado para escalas Likert; modela melhor dados não contínuos
<i>Weighted Least Squares Mean and Variance Adjusted (WLSMV)</i>	Variáveis ordinais com violação de normalidade	Ajusta média e variância do χ^2 ; mais robusto para dados ordinais; funciona melhor com amostras moderadas que o WLS tradicional
<i>Diagonally Weighted Least Squares (DWLS)</i>	Variáveis ordinais com violação de normalidade	Utiliza apenas a diagonal da matriz de pesos, reduzindo exigências computacionais; bom desempenho com correlações policóricas
<i>Unweighted Least Squares (ULS)</i>	Variáveis ordinais com violação de normalidade	Menor exigência computacional; adequado para correlações policóricas; pode funcionar bem com amostras moderadas

Nota. As descrições são baseadas em recomendações comuns na literatura sobre modelagem de equações estruturais (Bollen, 1989; Brown, 2015; Kline, 2016; Jöreskog & Sörbom, 1996; Muthén & Muthén, 2017).

- (ii) **Matriz de correlações (*Correlation Matrix*)** – A escolha da matriz de correlações a ser utilizada também constitui decisão metodológica fundamentada nas características da medida e na natureza dos dados. A AFC modela a matriz de associações observadas, seja de covariâncias ou correlações, estimando parâmetros que minimizem a discrepância entre a matriz empírica e a matriz implicada pelo modelo (Bollen, 1986; Brown, 2015). Dessa forma, o tipo e a qualidade da matriz empregada representam elementos centrais para a adequação das inferências derivadas da análise. O relato explícito da matriz utilizada é essencial, dado seu impacto direto nas estimativas de cargas fatoriais, nos índices globais de ajuste e na recuperação da estrutura latente, especialmente quando os itens são ordinais. Estudos de simulação indicam que, em dados do tipo Likert, o uso de matrizes policóricas tende a produzir estimativas mais precisas e ajuste mais adequado do que o emprego de correlações de Pearson, quando os pressupostos estatísticos correspondentes são atendidos (Flora & Curran, 2004; Holgado-Tello et al., 2008; Marôco, 2024). Fornecemos aqui uma breve descrição de algumas matrizes de correlações mais utilizadas na AFC (ver Tabela 2). Descrições detalhadas sobre as matrizes de correlações comumente usadas na AFC podem ser encontradas em Brown (2015), Kline (2016) e Muthén e Muthén (2017).

Tabela 2

Matrizes de correlação mais comuns em AFC

Matriz de correlações	Tipo de dado recomendado	Características
<i>Pearson Correlation Matrix</i>	Variáveis contínuas; Sensível a não normalidade dos dados	Correlação linear baseada em covariâncias observadas; padrão em AFC com dados contínuos
<i>Covariance Matrix</i>	Variáveis contínuas	Preserva variâncias originais; a escala de resposta dos itens influencia os parâmetros estimados.
<i>Polychoric Correlation Matrix</i>	Variáveis ordinais (e.g.: escalas Likert)	Escolha ideal para escalas Likert de 2 a 5 pontos. Se a escala tiver 6 ou mais categorias e a distribuição for normal, a Pearson começa a se aproximar da policórica (ver: Rhemtulla et al., 2012).
<i>Tetrachoric Correlation Matrix</i>	Variáveis dicotômicas	Assim como a matriz policórica exige amostras maiores do que as da matriz de Pearson para ser estável. Em amostras pequenas ($N < 200$), os índices de ajuste podem se tornar inflados.
<i>Polyserial Correlation Matrix</i>	Variável contínua + ordinal	É utilizada para estimar a relação entre variáveis contínuas e variáveis ordinais.

Nota. As descrições são baseadas em recomendações sobre AFC e estudos de simulações (Flora & Curran, 2004; Olsson et al., 1982; Rhemtulla et al., 2012)

- (iii) **Índices de ajuste absoluto (*Absolute Fit Indices*)** – Os índices de ajuste absoluto refletem as decisões analíticas previamente adotadas, bem como as características da medida e dos dados analisados. Esses indicadores permitem avaliar em que medida o modelo teórico especificado reproduz adequadamente a matriz de associações observada na amostra (Harrington, 2008). Entre os principais índices de ajuste absoluto destacam-se: o teste do qui-quadrado (*Chi-Square* – χ^2) e a razão qui-quadrado/graus de liberdade (χ^2/gf), o *Root Mean Square Error of Approximation* (RMSEA), o *Goodness-of-Fit Index*

(GFI), o *Adjusted Goodness-of-Fit Index* (AGFI), o *Root Mean Square Residual* (RMR) e o *Standardized Root Mean Square Residual* (SRMR) (Brown, 2015; Jackson et al., 2009).

a. O teste Qui-quadrado (*Chi-Square* – χ^2) e a razão Qui-quadrado/graus de liberdade (χ^2/gl) – O χ^2 é um teste de significância que avalia o ajuste global do modelo, examinando a discrepância entre a matriz de covariâncias observada e a matriz de covariâncias estimada pelo modelo (Hu & Bentler, 1999). Esse teste pressupõe normalidade multivariada; assim, desvios severos desse pressuposto podem resultar na rejeição do modelo mesmo quando ele está corretamente especificado (McIntosh, 2007). Idealmente, o χ^2 deve ser não significativo, o que significa que o modelo dos dados pode ser considerado um “bom” representante dos dados reais (Kline, 2016). Contudo, em amostras grandes ($N > 400$), o χ^2 quase sempre será significativo, mesmo que o modelo seja excelente. Por este motivo, a razão χ^2/gl , é utilizada como uma forma de relativização do valor do χ^2 , com o intuito de minimizar o impacto do tamanho da amostra sobre o teste (Tabachnick & Fidell, 2007; Wheaton et al., 1977).

b. RMSEA – é um indicador de ajuste que estima o grau de discrepância por grau de liberdade do modelo na população, avaliando o quanto o modelo se aproxima da matriz de covariâncias populacional (Byrne, 1998). Diferentemente do teste χ^2 , o RMSEA incorpora uma penalização pela complexidade do modelo e é frequentemente interpretado como uma medida de erro de aproximação. De forma simplificada, o índice estima o quanto o modelo, com parâmetros estimados na amostra, apresentaria discrepância se fosse aplicado à população. Recomenda-se que o RMSEA seja reportado juntamente com seu intervalo de confiança, que fornece informações adicionais sobre a precisão da estimativa e a plausibilidade do ajuste do modelo na população (McQuitty, 2004).

c. GFI e AGFI – O GFI é um índice de ajuste que expressa a proporção da variância-covariância observada e explicada pelo modelo, indicando o quão adequadamente o modelo reproduz a matriz de covariâncias observada nos dados (Diamantopoulos & Siguaw, 2000). O AGFI, por sua vez, consiste em uma versão ajustada do GFI que incorpora os graus de liberdade do modelo, penalizando modelos mais complexos e favorecendo soluções mais parcimoniosas (Brown, 2015).

d. RMR e SRMR – O RMR corresponde à raiz quadrada da média dos resíduos ao quadrado, calculados a partir da diferença entre a matriz de covariâncias observada e a matriz estimada pelo modelo, constituindo uma medida da discrepância média entre o modelo hipotetizado e os dados (Brown, 2015; Kline, 2016). De forma complementar, o SRMR consiste na versão padronizada do RMR, baseada nos resíduos das covariâncias padronizadas (ou correlações), o que reduz a influência das escalas de mensuração e facilita a interpretação do índice (Kline, 2016).

Por fim, a Tabela 3 apresenta um resumo dos principais pontos de corte recomendados na literatura para esses índices de ajuste.

Tabela 3

Principais pontos de corte recomendados na literatura para os índices de ajuste da AFC

Índice de Ajuste	Resumo	Ponto de corte
χ^2	Testa a discrepância entre a matriz observada e a matriz estimada pelo modelo.	Idealmente não significativo ($p > 0,05$)
χ^2/gl	Ajuste relativo do χ^2 considerando a complexidade do modelo.	<2,0 ou 3,0 – Excelente ou bom <5,0 – Aceitável >5,0 modelo mal especificado
RMSEA	Estima o grau de discrepância por grau de liberdade do modelo na população.	<0,05 ou < 0,06 – Ajuste excelente Entre 0,06 e 0,08 – Ajuste adequado Entre 0,08 e 0,10 – Ajuste Médio > 0,10 Ajuste pobre
CFI	Expressa a proporção da variância-covariância observada que é explicada pelo modelo.	>0,95 – Ajuste bom/excelente

Índice de Ajuste	Resumo	Ponto de corte
		>0,90 – Ajuste aceitável
<i>AGFI</i>	Versão Ajustada e mais parcimoniosa do CFI, considerando a complexidade do modelo.	<p>≥0,90 – Ajuste bom/aceitável</p> <p>≥0,85 – Ajuste aceitável para modelos complexos</p> <p><0,80 – Modelo mal especificado</p>
<i>RMR</i>	Apresenta uma estimativa da discrepância média entre o modelo hipotetizado e os dados.	Quanto mais próximos de 0 melhor.
<i>SRMR</i>	É uma versão padronizada do RMR que reduz a influência das escalas de mensuração.	<p>≤0,05 – Ajuste excelente</p> <p>≤0,08 – Ajuste bom</p> <p>>0,10 – Ajuste ruim</p>

Recomendações para um relato transparente e proporcional ao escopo do estudo.

A recomendação de estimar e reportar AFCs em estudos que utilizam instrumentos psicométricos, mesmo quando o objetivo principal da pesquisa não é a avaliação de suas propriedades psicométricas, pode ser compreendida à luz do princípio da proporcionalidade metodológica. De acordo com esse princípio, o nível de detalhamento e aprofundamento das análises deve ser compatível com o escopo do estudo, o grau de inovação na aplicação do instrumento e a complexidade das inferências empíricas propostas. Assim, espera-se que evidências de validade baseadas na estrutura interna sejam apresentadas de forma suficiente para sustentar as interpretações dos resultados, sem que isso implique necessariamente a realização de análises psicométricas extensivas, típicas de estudos de validação.

Assim, em estudos cujo objetivo principal não é a análise de parâmetros psicométricos, não se espera a realização de análises mais extensas, como comparações entre modelos alternativos ou testes formais de invariância. Nesses casos, os resultados das AFCs podem ser apresentados de forma breve e objetiva na subseção do método destinada à descrição dos instrumentos. O propósito dessa recomendação não é transformar todo o manuscrito em um estudo aprofundado das propriedades psicométricas das medidas utilizadas, mas assegurar que as inferências produzidas estejam sustentadas por evidências adequadas de validade baseadas na estrutura interna da própria amostra investigada.

Nesse contexto, é importante explicitar que a impossibilidade de testar empiricamente a estrutura da medida psicológica, por ausência de análise ou por tamanho amostral insuficiente para estimá-la adequadamente, constitui uma fragilidade metodológica. Estudos que propõem inferências complexas a partir de construtos latentes precisam assegurar que dispõem de condições estatísticas mínimas para sustentar tais inferências. Quando o delineamento ou o tamanho da amostra não permite avaliar a estrutura interna da medida, as interpretações derivadas dos escores tornam-se necessariamente mais limitadas e devem ser reconhecidas como tal. Naturalmente, estudos com tamanho amostral limitado, especialmente quando conduzidos em populações muito específicas ou de difícil acesso, devem ser analisados considerando as particularidades do contexto empírico e o potencial científico e inovador de seus achados.

Por outro lado, embora apresente os pontos de corte recomendados na literatura para os índices de ajuste da AFC, o presente artigo não os considera rígidos. Conforme a literatura metodológica tem demonstrado, critérios universais e descontextualizados podem conduzir a decisões simplificadas e, por vezes, equivocadas (Groskurth et al., 2024). Desse modo, a avaliação do ajuste de um modelo deve considerar múltiplos indicadores, como as características da amostra, o método de estimação utilizado, a complexidade do modelo e o arcabouço teórico que o fundamenta. A adoção de valores fixos como critério absoluto poderia levar a decisões mecânicas em detrimento de julgamentos metodologicamente fundamentados. Assim, para decisões metodológicas coerentes e fundamentadas, recomendamos a consulta a estudos de simulação e guias sobre a AFC (e.g.: Brown, 2015; Hooper et al., 2008; Jackson et al., 2009; Hu & Bentler, 1999).

Implicações para a pesquisa em Psicologia Organizacional e do Trabalho

As recomendações discutidas ao longo deste artigo têm implicações relevantes para o desenvolvimento metodológico da pesquisa em POT. Em um campo científico no qual grande parte dos fenômenos investigados é operacionalizada por meio de construtos latentes, como engajamento, comprometimento organizacional, liderança, bem-estar no trabalho e desempenho, a qualidade das inferências empíricas depende diretamente da adequação do modelo de medida utilizado para representar tais construtos.

Primeiramente, a apresentação de evidências de validade baseadas na estrutura interna, por meio de AFCs, contribui para fortalecer a replicabilidade das pesquisas na área. Estudos que reportam explicitamente a adequação do modelo de medida permitem que pesquisadores, avaliem de forma mais transparente se a estrutura teórica das escalas utilizadas é sustentada em diferentes contextos empíricos, populações e delineamentos de pesquisa (Bollen, 1989; Brown, 2015; Kline, 2016). Dessa forma, a prática de reportar evidências de validade baseadas na estrutura interna favorece não apenas a avaliação crítica dos resultados apresentados, mas também a cumulatividade do conhecimento científico (MacCallum & Austin, 2000; Vandenberg & Lance, 2000; AERA et al., 2014).

Em segundo lugar, a consideração explícita do modelo de medida tem implicações diretas para a validade das inferências empíricas produzidas (Bollen, 1989; Brown, 2015; Kline, 2016). Quando instrumentos psicométricos são utilizados sem a verificação de sua estrutura interna na amostra investigada, assume-se implicitamente que os indicadores observados representam adequadamente os construtos teóricos de interesse. Entretanto, a literatura psicométrica destaca que a validade das interpretações derivadas dos escores depende da acumulação de diferentes evidências empíricas, incluindo aquelas baseadas na estrutura interna das medidas (Cronbach & Meehl, 1955; Messick, 1995; AERA et al., 2014). Ademais, evidências acumuladas indicam que a estrutura fatorial de instrumentos pode variar quando aplicados em contextos culturais, organizacionais ou ocupacionais distintos daqueles em que foram originalmente desenvolvidos (MacCallum & Austin, 2000; Vandenberg & Lance, 2000).

Outra implicação importante diz respeito ao erro de medida (*measurement error*). Modelos estatísticos baseados exclusivamente em variáveis observadas, como regressões lineares tradicionais, assumem implicitamente que as variáveis utilizadas são livres de erro, o que raramente corresponde à realidade das medidas psicológicas (Bollen, 1989; Kenny, 2015; Kline, 2016). Ao estimar explicitamente a relação entre indicadores observados e construtos latentes, modelos baseados em equações estruturais permitem separar a variância verdadeira da variância associada ao erro de medida (Bollen, 1989; Brown, 2015). Como consequência, tais modelos tendem a produzir estimativas mais precisas das relações entre os construtos investigados (Bentler, 1990; Kline, 2016).

Além disso, a consideração explícita do modelo de medida torna-se particularmente relevante em estudos que investigam processos de mediação e moderação, cada vez mais frequentes na literatura em POT. Quando tais análises são conduzidas com variáveis observadas derivadas de escores compostos ou de médias de escalas, o erro de medida pode distorcer a magnitude das relações estimadas, levando a interpretações potencialmente equivocadas sobre os mecanismos psicológicos investigados (Bollen, 1989; Brown, 2015; Kline, 2016). A incorporação de variáveis latentes nesses modelos permite estimar com mais precisão as relações estruturais entre os construtos, aumentando a robustez das conclusões teóricas derivadas dos resultados empíricos (MacKinnon, 2008; MacCallum & Austin, 2000).

Por fim, a adoção de práticas mais transparentes no relato de modelos de medida contribui para o amadurecimento metodológico da área, estimulando uma integração mais consistente entre teoria, operacionalização empírica e análise estatística. Em um campo científico fortemente dependente da mensuração de construtos latentes, o fortalecimento do modelo de medida não é apenas um aprimoramento técnico, mas também um elemento central para a produção de evidências científicas mais robustas e cumulativas.

Considerações finais

As recomendações aqui apresentadas não devem ser compreendidas como meras exigências técnicas adicionais, mas como parte de um movimento mais amplo de fortalecimento das boas práticas na utilização de medidas em psicologia. No campo da POT, em que grande parte do conhecimento empírico é construída a partir da investigação de construtos latentes, a robustez do modelo de medida não constitui um elemento periférico, mas uma condição estrutural para a validade das inferências produzidas.

A recomendação de reportar índices de ajuste provenientes de AFCs, mesmo em estudos cujo foco principal não seja a investigação de parâmetros psicométricos, deve ser compreendida como parte de um movimento mais amplo de fortalecimento da qualidade metodológica, da transparência analítica e da solidez das evidências científicas produzidas na área. Tal recomendação não pretende transformar manuscritos em estudos psicométricos extensivos, tampouco impor padronizações rígidas e descontextualizadas. Ao contrário, fundamenta-se no princípio da proporcionalidade metodológica, segundo o qual o nível de detalhamento das análises deve ser compatível com o escopo do estudo e com a complexidade das inferências propostas.

A consolidação da POT como campo científico exige que seus achados estejam ancorados em medidas teoricamente consistentes e empiricamente sustentadas nas amostras investigadas. Nesse sentido, assegurar evidências de validade baseadas na estrutura interna das medidas psicológicas utilizadas representa um passo essencial para fortalecer a cumulatividade do conhecimento, reduzir ambiguidades interpretativas e ampliar a credibilidade nacional e internacional da produção científica da área.

Espera-se, assim, que estas recomendações contribuam para fortalecer os padrões de rigor metodológico nas pesquisas da área e para fomentar uma cultura científica cada vez mais comprometida com a integração entre teoria, método e evidência empírica. Mais do que qualificar relatos individuais, a adoção de práticas analíticas transparentes e fundamentadas pode contribuir para o amadurecimento metodológico do campo como um todo.

Referências

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7a ed.). APA.
- Barrett, P. (2007). Structural equation modeling: Adjudging model fit. *Personality and Individual Differences*, 42, 815–824. <https://doi.org/10.1016/j.paid.2006.09.024>
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Bentler, P. M. (2007). On tests and indices for evaluating structural models. *Personality and Individual Differences*, 42, 825–829. <https://doi.org/10.1016/j.paid.2006.09.020>
- Bollen, K. A. (1986). Sample size and Bentler and Bonett's nonnormed fit index. *Psychometrika*, 51, 375–377. <https://doi.org/10.1007/BF02294825>
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.
- Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling*, 7, 461–483. https://doi.org/10.1207/S15328007SEM0704_2
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. <https://doi.org/10.1037/h0040957>
- Diamantopoulos, A., & Siguaw, J. A. (2000). *Introducing LISREL*. Sage Publications.
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.
- Doğan, M. (2015). Influence of sample size, estimation method and normality on fit indices in confirmatory factor analysis. *Journal of Social Science Research*, 9(2), 1822–1833. <https://doi.org/10.24297/jssr.v9i2.4939>
- Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 56–83. <https://doi.org/10.1080/10705519909540119>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466–491. <https://doi.org/10.1037/1082-989X.9.4.466>
- Groskurth, K., Bluemke, M. & Lechner, C.M. (2024). Why we need to abandon fixed cutoffs for goodness-of-fit indices: An extensive simulation and possible solutions. *Behav Res* 56, 3891–3914. <https://doi.org/10.3758/s13428-023-02193-3>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed.). Pearson.
- Harrington, D. (2008). Assessing confirmatory factor analysis: Model fit and model revision. In *Confirmatory factor analysis* (online ed.). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195339888.003.0004>
- Holgado-Tello, F. P., Chacón-MoscOSO, S., Barbero-García, I., & Vila-Abad, E. (2008). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity: International Journal of Methodology*, 44(1), 153–166. <https://doi.org/10.1007/s11135-008-9190-y>
- Hooper, D., & Coughlan, J. & M. (2008). Structural equation modelling: Guidelines for determining model fit. In *The Electronic Journal of Business Research Methods*, 6, 53-60.
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jackson, D. L., Gillaspay, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, 14(1), 6–23. <https://doi.org/10.1037/a0014694>
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Scientific Software International.
- Kenny, D. A. (2015). *Measuring model fit*. Retrieved from <http://davidakenny.net/cm/fit.htm>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, 51, 201–226. <https://doi.org/10.1146/annurev.psych.51.1.201>
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Lawrence Erlbaum.
- Marôco, J. (2024). Factor analysis of ordinal items: Old questions, modern solutions? *Stats*, 7(3), 984–1001. <https://doi.org/10.3390/stats7030060>
- McDonald, R. P., & Ho, M.-H. R. (2002). Principles and practice in reporting statistical equation analyses. *Psychological Methods*, 7(1), 64–82. <https://doi.org/10.1037/1082-989X.7.1.64>

- McIntosh, C. N. (2007). Rethinking fit assessment in structural equation modelling: A commentary and elaboration on Barrett (2007). *Personality and Individual Differences*, 42(5), 859–867. <https://doi.org/10.1016/j.paid.2006.09.020>
- McQuitty, S. (2004). Statistical power and structural equation models in business research. *Journal of Business Research*, 57(2), 175–183. [https://doi.org/10.1016/S0148-2963\(02\)00276-5](https://doi.org/10.1016/S0148-2963(02)00276-5)
- Olsson, U., Drasgow, F., & Dorans, N. J. (1982). *The polyserial correlation coefficient*. *Psychometrika*, 47(3), 337–347. <https://doi.org/10.1007/BF02294164>
- Powell, D. A., & Schafer, W. D. (2001). The robustness of the likelihood ratio chi-square test for structural equation models: A meta-analysis. *Journal of Educational and Behavioral Statistics*, 26, 105–132. <https://doi.org/10.3102/10769986026002105>
- Prince, M. (2008). Measurement validity in cross-cultural comparative research. *Epidemiologia e Psichiatria Sociale*, 17(3), 211–220. <https://doi.org/10.1017/S1121189X00001305>
- Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17(3), 354–373. <https://doi.org/10.1037/a0029315>
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson Education.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4–70. <https://doi.org/10.1177/109442810031002>
- Wheaton, B., Muthén, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, 8, 84–136. <https://doi.org/10.2307/270754>

Contribuições:

Germano Gabriel Lima Esteves: conceituação, investigação, redação do manuscrito original, redação - revisão e edição.

Disponibilização de dados:

Os dados da pesquisa estão disponíveis mediante solicitação ao autor correspondente.

Conflitos de interesse:

O autor declara a inexistência de conflitos de interesse na realização e na comunicação dessa pesquisa.

Recebido: 7 de abril de 2026

Revisado: 29 de abril de 2026

Aceito: 29 de abril de 2026

Publicado: 30 de abril de 2024