

A Modelagem de Equações Estruturais com mínimos quadrados parciais: uma técnica estatística para os estudos em defesa e segurança internacional

The Structural Equation Modeling with partial least squares: a statistical technique for defense and international security studies

Resumo: A modelagem de equações estruturais com mínimos quadrados parciais (*Partial Least Squares Structural Equation Modeling* – PLS-SEM) é uma técnica estatística multivariada robusta, que se ajusta às pequenas amostras e permite aos pesquisadores responder uma série de perguntas inter-relacionadas de uma forma simples, sistemática e abrangente. Essa ferramenta foi criada especificamente para ser usada nas ciências sociais, em que nem sempre os estudos têm amplas amostras e novas teorias podem ser formadas a partir das constantes mudanças sociais. Consegue-se esse intento, portanto, ao modelar as relações entre múltiplos construtos dependentes e independentes, considerando diferentes tipos de medidas e diferentes variáveis. Este ensaio metodológico pretende apresentar a técnica estatística em questão, descrevendo as premissas do método, seus procedimentos, parâmetros de qualidade e limites. Ilustra-se esta análise com um exemplo aplicado por meio do uso do software SmartPLS. Dessa forma, este artigo apresenta a reflexão sobre a potencialidade dessa técnica estatística para pesquisas na área de segurança internacional e defesa.

Palavras-chave: estatística multivariada; predição; SmartPLS; construtos; modelos estruturais.

Abstract: Partial least squares structural equation modeling (PLS-SEM) is a robust multivariate statistical technique, which adjusts to small samples and allows researchers to answer a series of interrelated questions simply, systematically, and comprehensively. This tool was explicitly created to be used in the social sciences, in which studies do not always have large samples, and new theories can be formed from the constant social changes. It achieves this, thus, by modeling the relationships between multiple dependent and independent constructs, considering different types of measures and various variables. This methodological essay aims to present the statistical technique in question, describing the assumptions of the method, its procedures, quality parameters, and limits. This analysis is illustrated with an applied example by using the SmartPLS software. Thus, the article brings a reflection on the potential of this statistical technique for international security and defense research.

Keywords: multivariate statistics; prediction; SmartPLS; constructs; structural models

Angela Nogueira Neves 

Exército Brasileiro, Escola de Educação Física do Exército, Divisão de Pesquisa. Escola Superior de Guerra, Programa de Pós-Graduação em Segurança Internacional e Defesa. Universidade da Força Aérea, Programa de Pós-Graduação em Desempenho Humano Operacional. Rio de Janeiro, RJ, Brasil. angelanneves@gmail.com

Wanderson Roberto da Silva 

Universidade Estadual Paulista. Programa de Pós-Graduação em Alimentos, Nutrição e Engenharia de Alimentos Araraquara, SP, Brasil. Universidade Federal de Alfenas. Programa de Pós-Graduação em Nutrição e Longevidade Alfenas, MG, Brasil. wandersonroberto22@gmail.com

Recebido: 8 ago. 2022

Aprovado: 12 mar. 2023

COLEÇÃO MEIRA MATTOS

ISSN on-line 2316-4891 / ISSN print 2316-4833

<http://ebrevistas.eb.mil.br/index.php/RMM/index>



Creative Commons
Attribution Licence

1 INTRODUÇÃO

A modelagem de equações estruturais (*Structural Equation Modeling* – SEM) é uma técnica estatística de segunda geração, relativamente nova, que permite aos pesquisadores explicar as inter-relações entre múltiplos construtos (ou variáveis latentes) que foram medidos indiretamente por variáveis observáveis, ao combinar aspectos de análise fatorial e regressão múltipla (HAIR; *et al.*, 2014).

Diferencia-se das outras análises, como a regressão múltipla, análise multivariada de variância, análise discriminante e outras análises multivariadas por não focar em relacionamentos individuais entre variáveis dependentes e independentes, contudo, por considerá-las todas ao mesmo tempo. Essa característica possibilita a modelagem de equações estruturais atingir seu principal objetivo: expandir a capacidade do pesquisador para propor e confirmar novas teorias (HAIR; ANDERSON; BABIN, 2019).

Há dois tipos de SEM: a *Covariance-Based Structural Equation Modeling* (CB-SEM) e a *Partial Least Squares Structural Equation Modeling* (PLS-SEM). A CB-SEM foi a primeira a ser desenvolvida ao final da década de 1960 e início de 1970, baseada nos trabalhos de Jöreskog e se tornou popular por meio do software de análise por ele também criado, o sistema LISREL (MARÔCO, 2021). A CB-SEM é destinada a princípio a confirmar ou rejeitar teorias desenvolvidas a partir da verificação de relações sistemáticas entre múltiplas variáveis que foram medidas de forma empírica. A CB-SEM confirma a teoria proposta ao estimar uma nova matriz de covariância (a partir dos dados coletados em campo) que não é estatisticamente diferente da matriz de covariância original – estabelecida na proposição teórica do modelo de inter-relações. A CB-SEM emprega o estimador de Máxima Verossimilhança ou equivalentes para estimar as relações estabelecidas entre os construtos em uma matriz de covariância. Apesar de haver estimadores que compensam a distribuição não paramétrica dos dados – que se baseiam em equações policóricas – é premissa que os dados sejam aderentes à distribuição normal, e sejam coletados por mensuração minimamente intervalar. Ademais, a CB-SEM analisa modelos reflexivos, mas, aceita relações não recursivas (relação mútua e recíproca) e recursivas (em uma única direção) entre eles (HAIR *et al.*, 2014, 2019; HAIR; RINGLE; SARSTEDT, 2011; KLEM, 2006; SCHUMACKER; LOMAX, 2004).

Mas, se as suposições de CB-SEM não podem ser atendidas; ou objetivo da pesquisa é a previsão ao invés da confirmação de relações estruturais; ou modelos formativos estão previstos na teoria, então a PLS-SEM é o método recomendado. A PLS-SEM surge no final de década de 1970, nos trabalhos de Wold (professor orientador de Jöreskog), sobretudo, pensada para ser aplicada para as pesquisas nas ciências sociais, recebendo, algum tempo depois, os desenvolvimentos feitos por Lohmöller (1989). O método de estimação da PLS-SEM é o de mínimos quadrados ordinários (OLS), que pretende estimar as relações no modelo com o objetivo de minimizar a variância não explicada (erros) das variáveis dependentes (endógenas) no modelo. Seu algoritmo iterativo separadamente resolve os blocos do modelo de medida e, então, em uma segunda etapa, estima os coeficientes de caminho do modelo estrutural. Por conta dessa busca por ampliar a variância explicada da variável endógena, a PLS-SEM é a primeira técnica a ser considerada quando o pesquisador deseja desenvolver uma nova explicação, isto é, uma nova predição dos construtos para fundamentar uma nova teoria (HAIR *et al.* 2014; HAIR; RINGLE; SARSTEDT, 2011; VINZI *et al.*, 2010).

Na PLS-SEM, como o estimador OLS é livre de distribuição, os dados não necessariamente precisam ser aderentes à distribuição normal, nem aleatórios e/ou lineares, suportando bem desvios de assimetria e curtose. Os dados em análise podem ser medidos com escalas razão, intervalar (como dados recolhidos em escalas Stapel ou diferencial semântico), ordinal (como dados recolhidos em escalas tipo Likert) e dicotômicas (do tipo sim/não), esta última com alguma restrição. Ela trabalha tanto com modelos de medida formativos quanto reflexivos, mas unicamente com relações recursivas entre as variáveis latentes. Por não exigir aderência à distribuição normal, nem grandes amostras e nem dados lineares colhidos de forma aleatória, a PLS-SEM é chamada de *soft modeling* – que nada se refere a sua qualidade, poder e critério (VINZI *et al.*, 2010).

2 A PLS-SEM PODE SER ÚTIL NOS ESTUDOS EM DEFESA E SEGURANÇA INTERNACIONAL?

Pode-se dizer que os estudos de segurança internacional e defesa lidam com decisões tomadas nas esferas políticas em prol do equilíbrio de poder internacional e da manutenção da paz.

Essas decisões não são tomadas em um cenário simples, tampouco consideram poucas variáveis em sua elaboração, uma vez que são, sobretudo, decisões humanas. Logo, o processo de compreendê-las exige que seja considerada a multiplicidade de fatores que as causaram. Esses podem ser primariamente identificados e compreendidos por abordagens qualitativas na descrição dos fenômenos e na compreensão de suas dimensões. A PLS-SEM oferece um passo analítico complementar, pois permite prever, com dados empíricos, os efeitos das variáveis consideradas. A pesquisa com abordagem qualitativa permitirá dizer se há relação entre as variáveis, enquanto a abordagem quantitativa permitirá dizer o quanto elas podem prever umas às outras e se isso ocorre ao acaso ou não. Logo, o pesquisador poderá fazer previsões a respeito das mudanças de cenários, condições e atitudes, considerando que as decisões (variáveis endógenas) sempre dependem do contexto em que são tomadas.

Encontramos algumas pesquisas na área que se utilizaram dessa técnica estatística multivariada para investigar temas relacionados à segurança internacional e defesa, como cadeia logística (KOUVELIS; MUNSON, 2004; RAHIMI SHEIKH; SHARIFI; SHAHRIARI, 2017), diplomacia (AKBARIYE; VAZIFEDOUST; SALEH ARDESTANI, 2018), tráfico humano (RUDOLPH; SCHNEIDER, 2013), comportamento do eleitor (CWALINA; FALKOWSKI; NEWMAN, 2010), patriotismo e nacionalismo (KARASAWA, 2002) e estratégias globais (BOUQUET; BIRKINSHAW, 2011), indicando que a área tem aceitado essa técnica. Todavia, diante das demais alternativas metodológicas, parece que a SEM é ainda pouco explorada.

Este ensaio pretende apresentar ao pesquisador das Ciências Militares, Defesa e Segurança Internacional esse tipo específico de modelagem de equações estruturais, a PLS-SEM. Dessa forma, apresenta as premissas do método, seus procedimentos e parâmetros de qualidade, cujas informações foram sistematizadas a partir de livros e artigos clássicos e atualizados citados ao longo do texto. As informações foram organizadas de forma a conceituar seus elementos constituintes para depois apresentar as fases do processo estimativo – incluindo os parâmetros de ajuste, seguido pelas críticas feitas ao método. Para melhor apresentar a técnica, foi trazido um exemplo de uma análise feita no *software* SmartPLS.

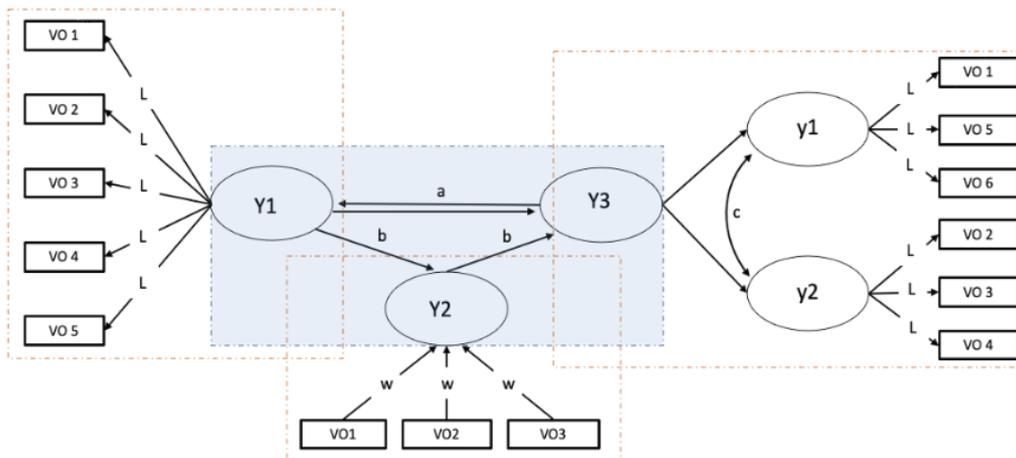
3 A TÉCNICA ESTATÍSTICA EM QUESTÃO: ELEMENTOS DA PLS-SEM

A PLS-SEM opera estimando inter-relações entre variáveis latentes (construtos). Variável latente é uma variável hipotética, que não pode ser medida diretamente, por exemplo, percepção de segurança. Entretanto, pode ser observada por indicadores selecionados, tais quais, os itens de um índice de percepção de segurança; que na PLS-SEM são chamados de variáveis observáveis. Os modelos PLS-SEM podem ser representados em gráficos e neles as variáveis latentes são, por convenção, expressadas por elipses, ao passo que, as variáveis observadas por retângulos (Figura 1).

As relações estabelecidas entre as variáveis determinam quais delas são dependentes e independentes do modelo. Na SEM, as variáveis independentes são chamadas de exógenas e as dependentes são chamadas de endógenas. A exógena é aquela que age como preditora em outra variável/construto no modelo teórico. É determinada fora do modelo e suas causas não são nele especificadas. A endógena é aquela que resulta de pelo menos uma relação de predição. O pesquisador conseguirá distinguir quais variáveis exógenas preveem cada variável endógena, apoiando-se na teoria e/ou em suas próprias experiências prévias (HAIR; ANDERSON; BABIN, 2019; HER-SHBERGER; MARCOULIDES; PARRAMORE, 2003; KLEM, 2006).

Há de se mencionar também que o modelo SEM é, na verdade, constituído por dois modelos: um estrutural e um de mensuração (Figura 1).

Figura 1 – Exemplo de um modelo de mensuração e estrutural



Fonte: Elaborado pelos autores, 2022

Os modelos de mensuração (*outer model*) estão em destaque nas três áreas hachuradas em laranja. Esses modelos versam sobre as relações entre as variáveis observáveis e as variáveis latentes. Eles podem conter variáveis latentes de primeira ordem, como Y1, no qual a variável latente é explicada por todas as variáveis observáveis em uma única dimensão. Podem conter também variáveis latentes de segunda ordem, nas quais uma variável latente (Y3) explica outras variáveis latentes de primeira ordem (y1 e y2), sendo que estas se conectam diretamente às variáveis observáveis, representando abstrações teóricas parciais para a formação do construto superior (Y3) (HAIR *et al.*, 2014; VINZI *et al.*, 2010)

É nesse modelo que se define a teoria da medida para cada variável latente, ou seja, se a variável latente será avaliada de forma reflexiva ou formativa. Graficamente, as setas partem da variável latente para as variáveis observáveis no modelo de mensuração reflexiva ($Y1$, $y1$ e $y2$). Nesse tipo de modelo, o construto afeta os indicadores, ou seja, é a partir da definição teórica que se busca elementos que podem ser observados. Cada variável latente afeta de forma distinta a variável observável, podendo ser medido pela carga fatorial, *outer loading*, (λ). As variáveis observáveis tendem a ser correlacionadas entre si e são, então, manifestações mais palpáveis das variáveis latentes, por exemplo, a capacidade de defesa internacional tem como manifestações a habilidade diplomática, qualidade das relações comerciais, papel no equilíbrio do poder e importância geopolítica.

Nos modelos de mensuração formativa, as variáveis observáveis são características da variável latente. Cada variável observável compõe uma parcela do construto, representada pelos pesos fatoriais, *outer weights* (ω), e uma pequena alteração em um deles altera o construto, como, a capacidade de combate se caracteriza pela quantidade de blindados, homens treinados, armas de destruição em massa, caças, munição disponível. Na representação gráfica (Figura 1) são vistas setas saindo dos indicadores (o equivalente às variáveis observáveis nos modelos formativos) para as variáveis manifestas ou domínios (o equivalente às variáveis latentes nos modelos formativos; $Y2$) (BECKER; KLEIN; WETZELS, 2012; HAIR *et al.*, 2014, 2019; VINZI *et al.*, 2010).

As relações entre os construtos em estudo estão previstas no modelo estrutural, ou *inner model*, que na Figura 1 está destacado na área azul. No modelo estrutural, o pesquisador define como as variáveis se relacionam. A força e a direção dessa relação são dadas pelo coeficiente estrutural (β). Graficamente, as relações são estabelecidas da esquerda para a direita, de forma que na margem esquerda da imagem estejam as variáveis exógenas ($Y1$), que são aquelas que apenas emitem setas, ao passo que, na direita, as variáveis endógenas ($Y3$), que recebem as setas de caminho; e algumas que recebem e emitem setas, também chamadas de endógenas ($Y2$), mas que dependendo da abordagem, podem também ser variáveis moderadoras (HAIR; RINGLE; SARSTEDT, 2011; VINZI *et al.*, 2010).

As setas indicam o tipo de relação entre as variáveis latentes e observáveis. Como se observa na Figura 1, se a seta for unidirecional (b), ela indicará uma relação recursiva entre as variáveis. Se a seta for bidirecional (a), ela indicará uma relação não recursiva entre as variáveis, ou seja, elas têm uma relação mútua e recíproca. A linha curvada, bidirecional (c) representa a correlação entre as variáveis latentes no modelo (GARVER; MENTZER, 1999; HAIR; ANDERSON; BABIN, 2019; HERSHBERGER; MARCOULIDES; PARRAMORE, 2003).

É no momento de definição do modelo estrutural que o pesquisador, de fato, delimita a sua proposição teórica e se propõe a responder perguntas como: quais explicações para as relações dos fatores sob estudo? quem os influencia significativamente? qual é a predição válida? Estas perguntas devem encontrar apoio no referencial teórico que suporta a investigação – mesmo que ele não esteja ainda definido por completo. Vale destacar que toda análise matemática feita nos dados com esse método estatístico será em vão se não houver apoio teórico. Não se deve perder de vista esse pressuposto.

4 O PROCESSO DE ESTIMAÇÃO DA PLS-SEM

As abordagens mais atuais propõem que o processo de decisão para estimação da PLS-SEM se dá por seis estágios (HAIR; ANDERSON; BABIN, 2019), que podem ser divididos em três fases: (i) o preparo da pesquisa; (ii) análise e ajuste dos modelos; e (iii) exploração adicional dos resultados.

Salienta-se que há hoje disponível ao pesquisador uma variedade de alternativas para tratar dados e realizar as análises da PLS-SEM com apoio de *softwares* (por exemplo, o SmartPLS; MPlus; PLS-Graph e o pacote PLS para o R). Algumas fórmulas apresentadas na seção foram feitas para facilitar o entendimento da definição do conceito. Em síntese, o trabalho do pesquisador é escolher suas medidas, determinar suas previsões, colher dados, programar o *software*, interpretar seus resultados para os modelos de mensuração estrutural, e, ainda, explorar mais os dados. Reflete, portanto, os seis estágios que serão descritos na sequência deste artigo, com exceção do domínio do software (programação), cuja escolha é a critério de cada pesquisador.

4.1 Fase 1: a preparação da pesquisa

O primeiro estágio trata sobre a definição dos objetivos de pesquisa e seleção dos construtos para o modelo. Quanto ao objetivo, PLS-SEM é o método estatístico multivariado a ser adotado quando se deseja investigar múltiplas relações em construtos para os casos nos quais a teoria subjacente não está bem definida – não há uma relação preditiva clara; ou quando se quer criar uma nova perspectiva teórica a partir das evidências empíricas geradas pela avaliação de construtos que podem estar relacionados; ou mesmo quando o problema de pesquisa lida com questões a respeito de como, por que e quando – já que a PLS-SEM aceita, basicamente, todos os níveis de mensuração (HAIR *et al.*, 2019).

Em seguida, devem-se selecionar os instrumentos adequados de avaliação para a geração de dados para a pesquisa. Instrumentos de medida com modelos reflexivos são avaliados por sua qualidade a partir de suas evidências métricas de validade e confiabilidade. Aqueles baseados em modelo formativos têm sua qualidade avaliada por seus pesos fatoriais e colinearidade entre os indicadores. Ter isso em perspectiva é de especial relevância nessa fase, afinal, testar hipóteses de predição entre construtos – que é o objetivo final da PLS-SEM – apenas se sustenta se a forma de medi-los oferecer dados confiáveis e avaliar adequadamente o construto em investigação.

As evidências de validade de uma medida indicam o que se está medindo (a representação do construto) e quais inferências podem ser feitas com base na mensuração (a interpretação do score) (URBINA, 2004). Elas devem ser consideradas no contexto da aplicação e, assim, as evidências de validade não são do instrumento em si, mas versam quanto à apropriação da medida em um determinado país, para uma determinada população (HURTZ; BANDEIRA; TRENTINI, 2015).

Classicamente podem-se distinguir até 31 tipos de validade (PASQUALI, 2007), entre os quais destacamos a validade de construto, que versa sobre a demonstração experimental de que o instrumento está medindo aquilo que ele se propõe a medir (BROWN, 2000). Há três subtipos de validade de construto: validade discriminante (evidências de que há fatores distintos na constituição do construto); validade convergente (evidências de que as variáveis selecionadas representam bem cada um dos fatores ou construto unifatorial); e validade nomológica (evidências de

associação entre construtos, entre variáveis observáveis e entre construtos e variáveis observáveis) (NUNNALLY; BERNSTEIN, 1994; PASQUALI, 2007; URBINA, 2004). Os testes estatísticos são diversos para verificar a validade de construto, mas, minimamente, a definição da estrutura fatorial – como os itens se organizam em fatores e como os fatores são identificados – é o necessário para considerar como evidência primária de validade de construto (MARÔCO, 2021).

Confiabilidade, de outra forma, refere-se às evidências de que o instrumento realiza medidas de forma consistente, livre de erros de medida (que vem a ser as flutuações nos escores que são irrelevantes ao que está sendo medido) (COZBY; BATES, 2001; URBINA, 2004). As escalas a serem escolhidas pelo pesquisador nesse primeiro estágio devem apresentar tanto evidências de validade quanto confiabilidade prévias.

Se o modelo for formativo, o pesquisador deve verificar a hipótese de evidências de ausência de multicolinearidade (correlações entre os indicadores) e se os pesos fatoriais (w) dos indicadores são relativa ou absolutamente significantes para a medida (HAIR *et al.*, 2017; ROBERTS; TATCHER, 2009).

O segundo estágio é o de desenho do estudo, no qual se define o tamanho amostral a ser coletado e o tratamento para dados faltantes. O cálculo do tamanho amostral deve considerar a preservação do poder estatístico do método – que é a capacidade do teste de identificar respostas corretas – o qual deve ser minimamente de 80% (COHEN, 1992). Nesse cálculo também se deve considerar a complexidade do modelo – o número de variáveis latentes, de variáveis observáveis e o número de relações causais estimadas – e a heterogeneidade amostral (HAIR; ANDERSON; BABIN, 2019; VINZI *et al.*, 2010).

Há algumas opções disponíveis para o cálculo do tamanho amostral e uma das primeiras regras de bolso para a definição do tamanho amostral mínimo requerido foi chamada de regra das 10x (BARCLAY; HIGGINS; THOMPSON, 1995), que determina multiplicação por 10 do número de indicadores do maior modelo formativo ou a multiplicação por 10 do maior número de setas (caminhos) direcionadas a um determinado construto no modelo. Essa regra foi motivo de sérias críticas (KOCK, 2018; KOCK; NADAYA, 2018) e hoje admite-se que é pouco – ou quase nada – adequada (KOCK, 2018; KOCK; NADAYA, 2018). Alternativamente, é recomendado que o tamanho amostral mínimo seja calculado por programas como *G*Power*¹ (utilizando-se de seus cálculos para testes de regressão); ou por métodos alternativos, tais quais, o de raiz quadrada inversa (que usa a raiz quadrada inversa do tamanho de uma amostra para a estimativa do erro padrão), e o método exponencial gama (que se baseia nas correções da função de suavização gama e exponencial aplicados ao método anterior) (KOCK; NADAYA, 2018). Além disso, o tamanho amostral pode mensurado por outras regras de bolso que consideram o nível de significância e a variável explicada (COHEN, 1992). O que o pesquisador não deve perder de vista é que, quanto mais heterogênea é sua população, maior número amostral é requerido, para apropriadamente avaliar as idiosincrasias das manifestações avaliadas (HAIR; RINGLE; SARSTEDT, 2011; HAIR *et al.*, 2012; HAIR; ANDERSON; BABIN, 2019; VINZI *et al.*, 2010).

Os dados faltantes devem ser considerados com cautela, pois o PLS-SEM é bastante sensível a eles. A forma mais criteriosa de lidar com essa questão é escolher a deleção *listwise*, isto é, excluir do banco de dados o sujeito (seja ele um Estado ou um indivíduo) para o qual não se tem a informação

1 Programa gratuito, disponível em: <http://www.psych.uni-duesseldorf.de/abteilungen/aap/gpower3/>. Sua vantagem é considerar o tamanho de efeito e o poder amostral para estimar o número amostral

de alguma das variáveis do estudo. A deleção *listwise* potencialmente diminui o banco de dados, e por isso, pode nem sempre ser a escolha ideal – apesar de ser a mais criteriosa. Após a coleta de dados, o pesquisador deve adicionalmente avaliar o padrão de dados faltantes para identificar que alguma variável observável foi sistematicamente deixada de ser respondida. Se assim ocorrer, o pesquisador deve considerar a existência de um viés em suas medidas (HAIR; ANDERSON; BABIN, 2019).

No terceiro estágio, o pesquisador define mais claramente qual hipótese testará em sua pesquisa. Começa pela proposição teórica feita pelo pesquisador em virtude das relações entre os construtos em estudo – que é uma primeira definição do modelo estrutural. Três elementos são críticos nessa etapa: (i) o pesquisador deverá ter conhecimento do tema sob investigação para determinar quais variáveis são endógenas e exógenas; (ii) deverá indicar as dependências e relações preditivas ou causais previamente identificadas na literatura entre as variáveis latentes; por fim, (iii) empregar aqueles instrumentos de medida selecionados no primeiro estágio, com evidências pertinentes de qualidade métrica (sejam formativas ou reflexivas) para dados coletados no país em que se conduz o estudo.

É nesse momento do planejamento da análise que o pesquisador define sua teoria estrutural e seus caminhos de predição. Uma revisão de literatura a respeito dos construtos envolvidos no estudo com evidências sobre a amostra estudada ampara essas decisões, assim como pesquisas anteriormente conduzidas pelo pesquisador ou seus parceiros de pesquisa (HAIR; ANDERSON; BABIN, 2019; VINZI *et al.*, 2010). É imprescindível que um amparo teórico seja constituído para suportar as predições.

Estando claro ao pesquisador que o método de análise estatística PLS-SEM atende aos objetivos da pesquisa, com quais medidas os dados serão coletados, em qual amostra e sob qual suposição teórica, o pesquisador deve ir a campo. Os estágios seguintes já consideram que o pesquisador tem dados coletados e está testando o ajuste métrico dos instrumentos à mostra e explorando a significância das relações causais propostas.

4.2 Fase 2: análise e ajustes dos modelos de mensuração e estrutural

O quarto estágio é o de avaliação dos modelos de mensuração. Esse processo é distinto caso os instrumentos de avaliação selecionados sejam reflexivos ou formativos. Quanto aos instrumentos reflexivos, trata-se de verificar as evidências de validade e confiabilidade das medidas para a amostra avaliada neste estágio. O pesquisador deve primeiro observar as cargas fatoriais das variáveis observáveis. Idealmente, as cargas fatoriais devem ser iguais ou superiores a 0,71, o que indicaria aproximadamente 50% da variação da variável latente. Entretanto, valores próximos são tolerados, desde que não comprometam a confiabilidade ou validade; mas valores abaixo de 0,40 têm alta recomendação de eliminação. Algumas variáveis observáveis podem ser eliminadas do modelo de mensuração em prol de melhorar a qualidade da medida, e isso não é incomum. Quando o pesquisador eliminar todas as variáveis observáveis com cargas fatoriais muito baixas ($< 0,40$), ele deve continuar o processo de refinamento da medida – inclusive levando em conta outras eliminações – considerando o impacto na confiabilidade e validade da medida (HAIR *et al.*, 2009, 2014, 2019).

O tipo de confiabilidade no âmbito da PLS-SEM avaliada é a consistência interna, isto é, a coerência interna dos indicadores em relação à variável latente. A PLS-SEM permite avaliar a confiabilidade de construto, que indica o “grau em que um conjunto de indicadores de construtos

latentes é consistente em suas mensurações” (HAIR *et al.*, 2009, p. 467). Pode-se usar a seguinte fórmula para calcular a confiabilidade de construto:

$$\text{confiabilidade de construto} = \frac{(\sum_{i=1}^n \lambda_i)^2}{(\sum_{i=1}^n \lambda_i)^2 + (\sum_{i=1}^n e_i)}$$

Em que Σ é somatório, λ_i é a carga fatorial padronizada das variáveis observáveis da variável latente e e_i é o erro de mensuração de cada item da variável latente, calculado como 1 - confiabilidade da variável observável.

Adicionalmente, pode-se avaliar também a consistência interna com outros testes e adicionar evidências de confiabilidade à medida. O teste clássico para gerar essa evidência é o teste alpha de *Cronbach*, que gera o coeficiente *alpha* de *Cronbach* (α) a partir das correlações entre as variáveis observáveis. Todavia, esse indicador tem sido bastante questionado como uma medida adequada às escalas não intervalares, ou mesmo se realmente representa a consistência interna de uma medida (SIJTSM, 2009). Há ainda defensores do seu uso e ele tende a ser mantido pela possibilidade de interpretação de evidências acumuladas (TAVAKOL; DENNICK, 2011). Como alternativa, o pesquisador pode empregar outros testes que corrigem algum desvio que o teste de alpha de Cronbach pode ter para as escalas não intervalares, como o teste *alpha* ordinal, *omega*, e o *greatest lower bound* (PETERS, 2014). Para todos os testes de confiabilidade mencionados, valores acima de 0,70 são recomendados, sendo tolerados os valores acima de 0,60 (HAIR; ANDERSON; BABIN, 2019).

A validade de construto é avaliada na sequência. O pesquisador deve gerar evidências de validade convergente e discriminante. A validade convergente é estabelecida quando a variância referente ao erro da medida é inferior à variância capturada, o que indica a extensão cujas variáveis observáveis convergem no construto. Essa é estabelecida com ajuda da variância média extraída (AVE) por meio da fórmula:

$$\text{Variância média extraída} = \frac{\sum_{i=1}^n \lambda_j^2}{n}$$

Em que Σ é somatório, λ_i é a carga fatorial padronizada das variáveis observáveis da variável latente e n é o número de itens em cada fator. Valores acima de 0,50 são considerados adequados (HAIR *et al.*, 2014, 2019).

A validade discriminante, por sua vez, aponta a extensão em que uma variável latente se diferencia de outra. Pode ser avaliada por intermédio da análise das correlações cruzadas, do critério de Fornell e Larcker e do teste HTMT (*heterotrait-monotrait ratio of correlations*). Quanto às análises das cargas fatoriais cruzadas, os indicadores de carregamento fatorial no construto atribuído devem ser maiores do que todos os carregamentos de outros construtos (HAIR; RINGLE; SARSTEDT, 2011).

O critério de Fornell e Larcker (FORNELL; LARCKER, 1981) compara a raiz quadrada da AVE com a correlação entre os construtos. A lógica que sustenta esse indicador é a de que uma variável latente deve explicar melhor a variância de seu próprio indicador, em vez da variância de outras variáveis latentes. Portanto, a raiz quadrada da AVE de cada construto deve ter um valor maior do que as correlações com outros construtos latentes (HAIR *et al.*, 2014).

A alternativa mais recente para investigar a validade discriminante é o teste HTMT. Henseler, Ringle e Sarstedt (2015) demonstraram o desempenho superior desse método em relação às duas alternativas anteriormente apresentadas, por meio do estudo de simulação de Monte Carlo, recomendando seu uso. Valores HTMT próximos a 1 indicam falta de validade discriminante. Alguns autores sugerem um limite de 0,85 (KLINE, 2015) e até mesmo de 0,90 (TEO; SRIVASTAVA; JIANG, 2008). Para Henseler, Ringle e Sarstedt (2015), os valores devem estar entre -1 e 1 para evidenciar validade discriminante. Menores valores provocam evidências mais importantes de validade discriminante.

O ajuste dos instrumentos de medida para atender os critérios de confiabilidade e validade se dá, sobretudo, pela inspeção das cargas fatoriais para eliminação dos itens com cargas muito baixas ($\lambda < 0,40$) e pela decisão de eliminar ou manter itens com cargas fatoriais entre 0,40 e 0,70. Itens que não impactam negativamente os indicadores de validade e confiabilidade com carga fatorial entre 0,69 e 0,41 podem ser mantidos, de forma a conservar melhor a validade de conteúdo da medida (HAIR; ANDERSON; BABIN, 2019).

Mas se há instrumentos com modelos de medidas formativos, o pesquisador deve estar ciente que a avaliação da qualidade do modelo de mensuração que contém estes instrumentos envolve outros parâmetros, pois aqueles até agora descritos baseiam-se largamente na correlação entre as variáveis observáveis. Logo, como se espera que os indicadores sejam independentes ou que, no máximo, sejam fracamente correlacionados nos modelos formativos, as medidas até o momento apontadas para evidenciar a validade e a confiabilidade não devem ser usadas nesse caso (ROBERTS; TATCHER, 2009). Por isso, as análises devem se focar em: (i) determinar se há correlações indesejadas entre dois ou mais indicadores – colinearidade; e (ii) determinar os pesos fatoriais (w) dos indicadores (HAIR *et al.*, 2017).

A colinearidade ocorre quando há correlação entre dois os indicadores – o que é problemático nesse modelo, pois se assume que cada indicador contribui de forma independente para o construto (em mais indicadores estando envolvidos, essa situação é chamada de multicolinearidade). Quando ocorre a colinearidade (ou a multicolinearidade), há aumento dos erros padrão e queda da capacidade de estimar os pesos fatoriais corretamente – tanto em seu valor quanto em sua significância. Essa é uma questão de particular preocupação para amostras pequenas, nos quais os erros padrão são, geralmente, grandes (HAIR *et al.*, 2014).

A inspeção da multicolinearidade pode ser avaliada pelo cálculo da tolerância (TOL). TOL versa sobre o quanto da variância de um indicador não foi explicada pelos outros indicadores na mesma variável latente (ou manifesta, preferencialmente neste caso). TOL é assim calculada:

$$TOL = R_{x1}^2$$

em que R_{x1}^2 = proporção da variância de x_i associada a outros indicadores. Há um valor de TOL para cada indicador do modelo e todos devem ser calculados. Uma medida relacionada a TOL é o fator de inflação da variância (VIF) que pode ser calculado como:

$$VIF = 1/TOL$$

Essas duas medidas de colinearidade podem ser usadas na análise como critério de retenção no instrumento, embora, informar somente a VIF tornou-se a prática mais comum.

São considerados como aceitável $VIF < 5$ e preferível $VIF < 3,3$ (HAIR *et al.*, 2017; DIAMATOPOULOS; SIGUAW, 2006). Para os indicadores que excedem esses valores, o pesquisador deve considerar eliminá-los do modelo, já que uma grande parcela de sua variância está sendo explicada pelos demais indicadores. Todavia, ele deve ter alguma certeza de que essa exclusão não afetará a constituição da variável manifesta (BIDO *et al.*, 2010; LATAN, RAMLI, 2013).

O peso fatorial do indicador (ω) trata sobre a contribuição de cada indicador à variável manifesta. Ele deriva da regressão múltipla do escore da variável manifesta (variável dependente) e dos indicadores (variáveis independentes). O valor de ω é padronizado, logo, pode ser comparado com os demais da mesma variável manifesta, permitindo ao pesquisador avaliar sua importância relativa na formação do construto (Hair *et al.*, 2014). Para ser importante, ω deve ser significativa e, preferencialmente, embora não necessariamente, $\omega > 0,50$. Quando ω é não significativa, porém tem carga fatorial (λ) alta (ou seja, acima de 0,50), o indicador deve ser interpretado como absolutamente importante (isto é, a informação que ele dá é importante, contudo, não considera os demais indicadores). Nessa situação, o indicador geralmente seria mantido no modelo formativo. Mas, quando um indicador tem ω não significativa e $\lambda < 0,50$, o pesquisador deve decidir se deve reter ou excluir o indicador examinando sua relevância teórica e a sobreposição de conteúdo potencial com outros indicadores do mesmo construto (HAIR *et al.*, 2014).

O quinto estágio se inicia quando o pesquisador pode passar ao exame do modelo estrutural, estando asseguradas as evidências de validade, confiabilidade e/ou colinearidade, valor e significância dos pesos fatoriais dos indicadores do modelo de mensuração. O primeiro passo do quinto estágio é avaliar a colinearidade entre os construtos exógenos (preditores). Calcula-se o fator de inflação de variância (VIF) para cada variável latente preditora e, por regra, maiores valores de VIF indicam colinearidade. O pesquisador também pode inferir a colinearidade pela execução de um teste de correlação bivariada entre as variáveis latentes predictoras, buscando por correlações inferiores a 0,50 para apontar ausência de colinearidade. Se houver indícios dessa ausência, o pesquisador pode pensar em criar um fator de segunda ordem para o construto em questão, a fim de lidar com o problema sem perder variáveis no modelo (HAIR; ANDERSON; BABIN, 2019).

Em seguida, o pesquisador pode ocupar-se das relações de predição em si (o modelo estrutural). O coeficiente de determinação (R^2) varia de 0 a 1, e maiores valores indicam maior capacidade de explicação do modelo estrutural – e, então, melhor predição das variáveis endógenas. Em termos gerais, valores de 0,25; 0,50 e 0,75 indicam efeitos fracos, moderados e substanciais, respectivamente. Mas na área de ciências sociais, os valores limites 0,02; 0,13 e 0,26 são recomendados para interpretação como fracos, moderados e substanciais, respectivamente (COHEN, 1988; HAIR *et al.*, 2014).

O tamanho de efeito (f^2) representa a alteração que o coeficiente de determinação sofre quando uma variável exógena é omitida no modelo, podendo ser interpretado como o quanto cada construto é útil para o ajuste do modelo. Valores de 0,02; 0,15 e 0,35 são considerados como efeitos pequenos, médios e grandes de uma variável latente exógena e valores menores que 0,02 indicam que não há efeito – nenhuma capacidade de predição (HAIR *et al.*, 2014; 2019).

O próximo passo é a avaliação do poder preditivo do modelo (Q^2 ; indicador de *Stone-Geisser*). Via de regra, o seu valor é obtido com um procedimento de *blindfolding* considerando-se a abordagem da redundância cruzada, disponibilizada na maioria dos softwares de análise da

PLS-SEM. Valores superiores ou iguais a zero para cada construto endógeno do modelo indicam que a acurácia preditiva do modelo é aceitável, sendo que maiores valores indicam maior acurácia preditiva (HAIR *et al.*, 2014, 2019; VINZI *et al.*, 2010).

Por fim, o pesquisador avalia o tamanho e a significância das relações preditivas, dadas pelos coeficientes estruturais (que graficamente são representados pelas setas no modelo e conceitualmente explicitam as relações de predição). Os coeficientes estruturais devem ser significantes – se $p < 0,05$, $t > 1,96$ e se o intervalo de confiança não contém zero. Caso sejam identificados valores maiores que esses, é um indicativo adicional de haver colinearidade no modelo (Tabela 1). Os valores dos coeficientes estruturais devem ser interpretados à luz da teoria para avaliar sua importância e sua interpretação matemática é tal qual os betas (β) das regressões lineares simples ou ordinárias. Valores que podem ser considerados aceitáveis dependem da complexidade do modelo e do contexto da pesquisa (HAIR; ANDERSON; BABIN, 2019; RINGLE; SILVA; BIDO, 2014).

Tabela 1 – Síntese dos ajustes do Smart PLS

INDICADOR	PROPÓSITO	VALORES REFERENCIAIS/CRITÉRIO	REFERÊNCIAS
VIF	Avaliação da multicolinearidade.	VIF < 5 VIF < 3,3 (maior rigor)	HAIR <i>et al.</i> (2017)
AVE	Avaliação da validade convergente.	AVE 0,50	HENSELER; RINGLE; SINKOVICS (2009)
Método HTMT	Avaliação da validade discriminante.	HTMT < 0,85	HAIR <i>et al.</i> (2019)
Confiabilidade composta	Avaliação da confiabilidade do modelo.	CC 0,70	HAIR <i>et al.</i> (2017)
Teste T de Student	Avaliação das significâncias das correlações e regressões.	$t > 1,96$	HAIR <i>et al.</i> (2017)
Coefficiente de determinação de Pearson (R^2)	Avaliam a porção da variância das variáveis endógenas, que é explicada pelo modelo estrutural.	Para a área de ciências sociais e comportamentais, $R^2 = 2\%$ seja classificado como efeito pequeno, $R^2 = 13\%$ como efeito médio e $R^2 = 26\%$ como efeito grande.	COHEN (1988)
Indicador de Cohen (f^2)	Avalia-se quanto cada construto é útil para o ajuste do modelo.	Valores de 0,02, 0,15 e 0,35 são considerados pequenos, médios e grandes.	HAIR <i>et al.</i> (2017)
Indicador de Stone-Geisser (Q^2)	Avalia a acurácia do modelo ajustado.	$Q^2 > 0$	HAIR <i>et al.</i> (2017)
Coefficiente estrutural (β)	Avaliação das relações causais.	Interpretação dos valores à luz da teoria.	HAIR <i>et al.</i> (2017)

Fonte: Ringo, Silva e Bido (2014); Hair *et al.*, 2017, 2019

4.3 Fase 3: exploração adicional dos resultados

No sexto estágio do processo de estimação da PLS-SEM, o pesquisador pode fazer análises adicionais para melhor explicar, especificar ou mesmo interpretar seu modelo. Das análises possíveis, destacamos dois: a investigação de heterogeneidade na amostra e a presença de variáveis mediadoras ou moderadoras.

Por meio de uma análise multigrupo, o pesquisador pode avaliar se há diferença entre subgrupos amostrais formados a partir de variáveis categóricas do estudo (por exemplo: continente, sistema de governo etc.) e esta é chamada de heterogeneidade observada. As subamostras as quais se deseja investigar a heterogeneidade devem atender aos critérios de poder e nível de significância amostral (ou seja, não devem ser menores que o mínimo calculado, considerando-se o modelo estudado). Serão calculados os coeficientes estruturais para cada amostra e será feita a análise da variação entre eles. Caso seja indicado que não há variação ($p > 0,05$), será concluída ausência de heterogeneidade amostral. A abordagem para essas análises é chamada de PLS-MGA (*PLS – multigroup analysis*) e está embutida nos softwares mais comuns.

Pode-se avaliar também a heterogeneidade não observada – os grupos se diferem por uma característica não observada pelo pesquisador. Ela pode ser abordada pela probabilidade dos participantes de amostra de pertencer a segmentos (subamostras) e pela diferença de coeficientes estruturais dos segmentos. É o pesquisador que deve prever o número de segmentos possíveis, estimado a partir da razão entre o número total amostral e o número mínimo amostral requerido. O número de segmentos será o menor valor inteiro mais próximo dessa razão. Essa abordagem é feita pelo método FIMIX (*Finite Mixture Partial Least Squares*) e ela é específica para calcular a heterogeneidade no modelo estrutural. Outra abordagem é o PLS-POS (*Prediction-Oriented Segmentation*) que também observa a heterogeneidade no modelo estrutural e de mensuração, mas com a vantagem que o faz tanto com modelos com variáveis formativas e reflexivas e gera informações acerca da variância explicada (R^2) ao formar grupos homogêneos e ao reatribuir observações, caso venha melhorar os critérios de qualidade dos modelos, além de trabalhar com medidas de distância específicas para variáveis formativas. Hair *et al* (2017) recomendam que esses dois métodos sejam usados em conjunto, começando pelo FIMIX e pela análise de segmentos. Se não rejeitar a existência de heterogeneidade no modelo estrutural, sugere continuar as análises com o PLS-POS. Por fim, a terceira opção é o REBUL-PLS (*Response-Based Procedure for Detecting Unit Segments*) que analisa a heterogeneidade no modelo estrutural e de mensuração com variáveis reflexivas, observando os resíduos dos modelos e das variáveis latentes (HAIR *et al.*, 2017, 2019).

Nos casos em que uma variável é predita por uma variável exógena e ao mesmo tempo prediz uma variável endógena (na Figura 1 é o caso da variável Y2), é válido investigar se ela pode ser uma variável que faz mediação ou que modera a predição da variável exógena sobre a variável endógena. Uma variável moderadora é aquela que contamina a relação preditiva entre a variável independente e dependente e deve então, ser controlada, pois afeta a força e a direção da predição (por exemplo, entre a necessidade de deter um inimigo e realmente fazê-lo, esse desfecho pode ser moderado pela confiança em sua habilidade de combate). Uma variável mediadora, por sua vez, é aquela que explica o processo de relação entre duas variáveis, ou seja, é um intermediário entre a variável independente e dependente (por exemplo, a qualidade do sono pode afetar o desempenho

em combate, em função do estado de alerta cognitivo). Saber se a variável tem o efeito mediador ou moderador permite avaliar os efeitos totais entre os construtos, permitindo uma avaliação mais completa da predição (HAIR *et al.*, 2017, 2019).

5 LIMITAÇÕES E CRÍTICAS À PLS-SEM

Apesar da PLS-SEM permitir lidar com dados de diferentes níveis de mensuração; dados não aderentes à distribuição normal; amostras relativamente pequenas; suporte teórico relativamente diminuto, colaborando para o amadurecimento de uma teoria em investigação, há críticas ao método.

Primeiro, critica-se que o modelo estrutural não permite que haja relações não recursivas (influência recíproca entre variáveis). Se a predição teórica do pesquisador ou evidências prévias indicarem que considerar essa relação é importante ao modelo, o pesquisador deve considerar não usar a PLS-SEM, mas a CB-SEM (HAIR *et al.*, 2014, 2017, 2019; VINZI *et al.*, 2010).

Segundo, afirma-se que as estimativas da PLS-SEM não são eficientes e potencialmente viesadas quando comparadas às medidas obtidas pelo método CB-SEM (por exemplo, MARCOULIDES; SAUNDERS, 2006; MCINTOSH; EDWARD; ANTONAKIS, 2014; RÖNKKÖ, 2014; RÖNKKÖ; EVERMAN, 2013). O argumento central que ampara essas críticas é que as estimativas PLS-SEM são inconsistentes por serem agregadas das variáveis observadas e incluírem erros de medição. Esse viés tende a se manifestar em estimativas um pouco mais altas para as cargas fatoriais e estimativas de coeficientes de caminho mais baixas.

Em resposta, os pesquisadores que trabalham com a PLS-SEM argumentam que já demonstrado que as estimativas se aproximarão dos valores dos parâmetros verdadeiros quando aumentarem o número de indicadores por construto e o tamanho da amostra (VINZI *et al.*, 2010). Para aqueles que comparam PLS-SEM e CB-SEM, já foi demonstrado que as diferenças matemáticas entre as estimativas são irrelevantes, sendo a PLS-SEM uma escolha tão boa quanto a CB-SEM para tratar modelo reflexivo e superior à esta última, para trabalhar com modelos formativos (SARSTEDT *et al.*, 2016).

6 EXEMPLO APLICADO

Para exemplificar a aplicação dessa técnica estatística em um banco de dados, escolheu-se usar o *software* SmartPLS. Ele é intuitivo, simples de ser usado, com licenças gratuitas para estudantes². Pode ser encontrado em livros, artigos e vídeos para dar suporte aos usuários, além de esclarecimentos técnicos sobre o *software*. Neste exemplo, foi usada a quarta versão do SmartPLS, lançada em agosto de 2022.

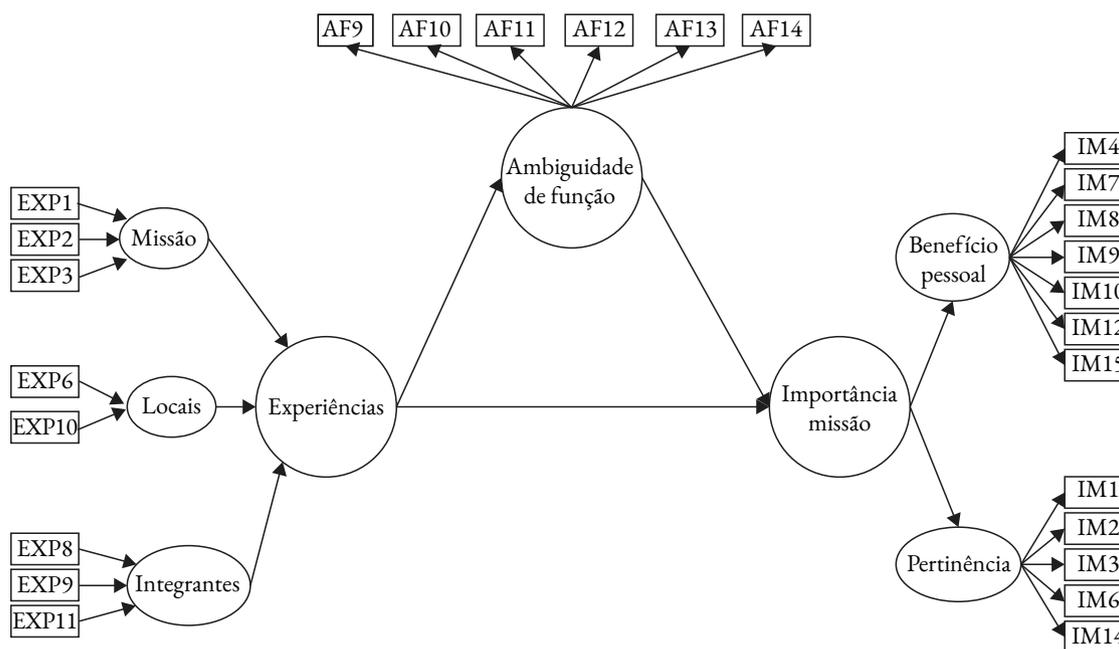
Foi testado um modelo hipotético cujas experiências do militar no terreno predizem a avaliação da importância de uma missão, podendo ou não ser mediada pela ambiguidade de função. Esse modelo foi teoricamente concebido a partir de evidências que margeiam o tema, reunidas

2 <https://www.smartpls.com/>

por revisão sistemática da literatura – uma maior discussão acerca das bases teóricas desse modelo usado para exemplo foge da atenção deste artigo, mas podem ser vistas em Neves (2022).

A primeira variável exógena, experiências no terreno, está operacionalizada em um modelo formativo. A segunda variável exógena e a variável endógena estão operacionalizadas em um modelo reflexivo. Tanto a experiência no terreno quanto a importância da missão são variáveis latentes de segunda ordem, ou seja, compostas por outras variáveis latentes (missão, locais e integrantes; pessoal e pertinência, respectivamente). Todas as medidas foram previamente analisadas em um estudo métrico que apontou quais variáveis observáveis deveriam ser retidas para cada variável latente, em uma amostra similar à do banco de dados aqui usado (Neves, 2022). Há uma predição direta entre a experiência no terreno e a importância da missão, e há a predição de que pode ser mediado pela ambiguidade de função. As elipses representam as variáveis latentes, os retângulos as variáveis observáveis e as setas os caminhos de predição (Figura 2).

Figura 2 – Modelo teórico hipotético para o exemplo aplicado



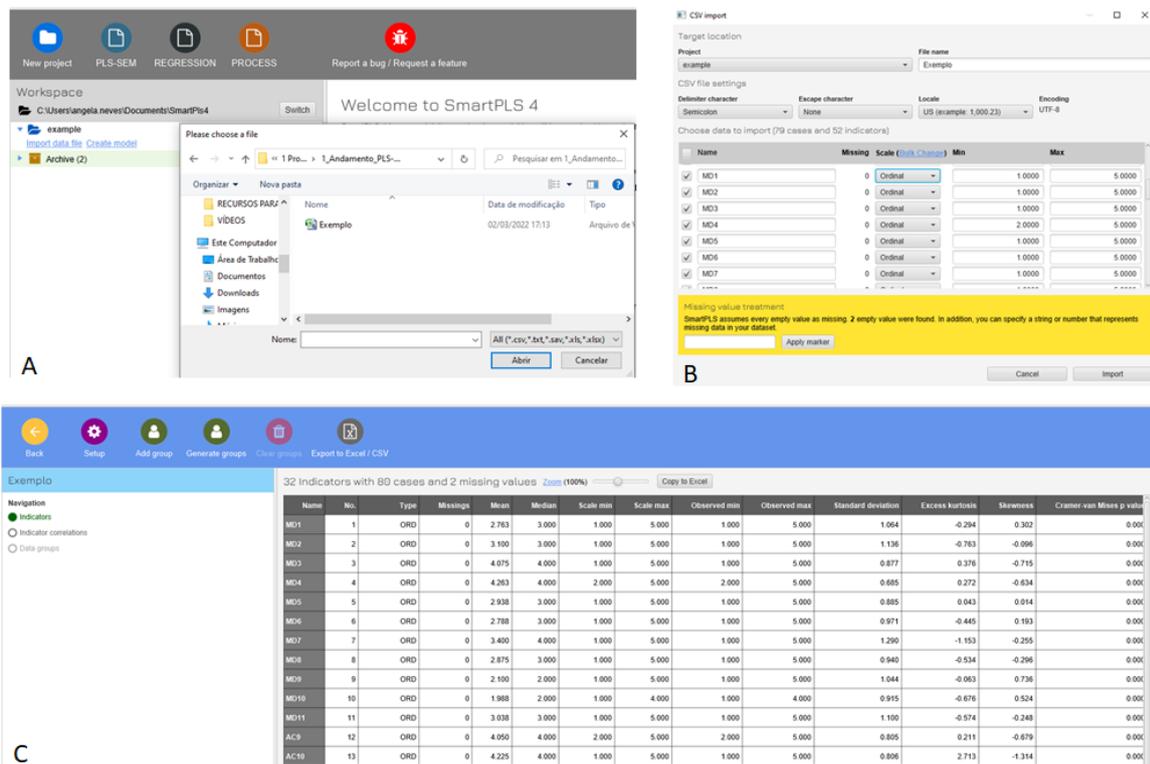
Fonte: Elaborado pelos autores, 2022

O tamanho amostral mínimo para testar esse modelo foi calculado no software G*Power, conforme as recomendações de Bido *et al* (2014). Logo, a variável endógena (importância da missão) é predita por duas variáveis exógenas (experiências no terreno e ambiguidade de função) e considerando-se tamanho de efeito de 0,15, 80% de poder e alpha de 0,05, o tamanho amostral mínimo sugerido é de 68 respondentes, sendo o ideal o triplo desse valor. Para esse exemplo, o banco de dados está constituído por 80 respondentes.

Os dados podem ser organizados em uma planilha do software de preferência do pesquisador, devendo ser obrigatoriamente salva no formato *.csv, sendo que na primeira linha da planilha são identificados os rótulos (nomes) das variáveis avaliadas.

Ao iniciar o programa, escolha a opção *new project* e o nomeie, de acordo com sua pesquisa. Em seguida, será aberta uma pasta do projeto no menu à esquerda com duas opções: *import data* e *create model* (Figura 3A). Comece pela primeira opção, escolhendo a planilha *.csv com os dados da pesquisa. Assim que os dados forem importados, o SmartPLS abrirá uma nova janela para que o pesquisador ajuste o nível de mensuração dos dados e informe se há algum código para dados faltantes (Figura 3B). Clique em *import* e a planilha de dados será carregada com a opção de identificar grupos amostrais, se for pertinente (Figura 3C). Uma vez importados corretamente, os dados serão identificados na pasta do projeto, na cor verde.

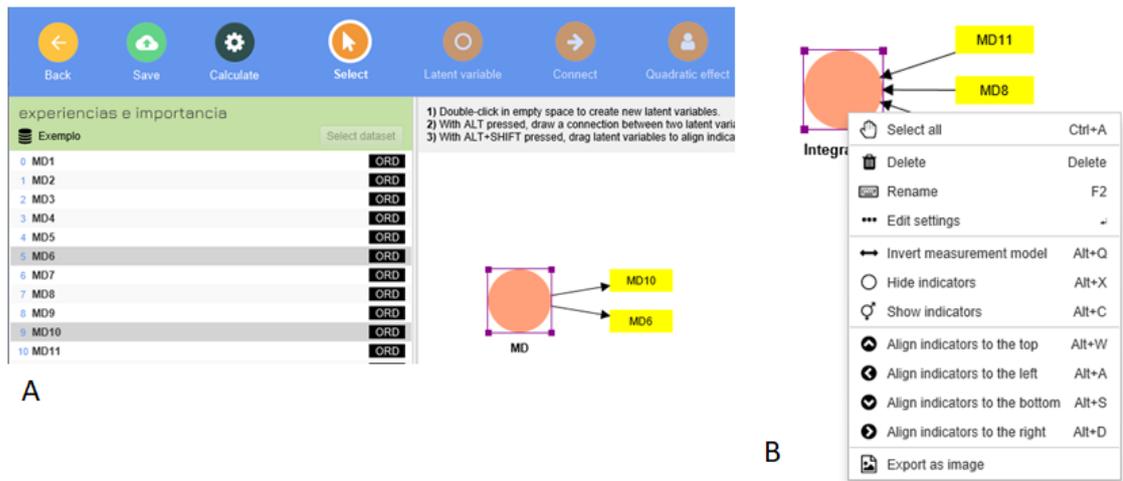
Figura 3 – Importação do banco de dados para o software



Fonte: Elaborado pelos autores, 2022

Agora escolha a segunda opção, *create model*. Será aberta uma janela para que modelo seja nomeado e para escolha do tipo de modelo. Para a análise de equações estruturais, escolha *PLS-SEM*. Clique em *Save*. No menu à esquerda, estarão identificadas todas as variáveis observáveis e indicadores de seu modelo. Selecione todas as que refletem ou formam a mesma variável latente (no exemplo, itens MD1, MD2, MD3 foram selecionados para a variável latente Missão) e as arraste para a direita (janela gráfica), já a posicionando de acordo com o modelo hipotético. Será aberta uma janela de texto para nomear corretamente a variável latente. Pressione *enter* (Figura 4A). Com o botão esquerdo do mouse pode ser aberto um menu para ajuste da variável latente, incluindo o tipo de modelo (se formativo ou reflexivo), o posicionamento das variáveis observáveis, entre outras opções (Figura 4B).

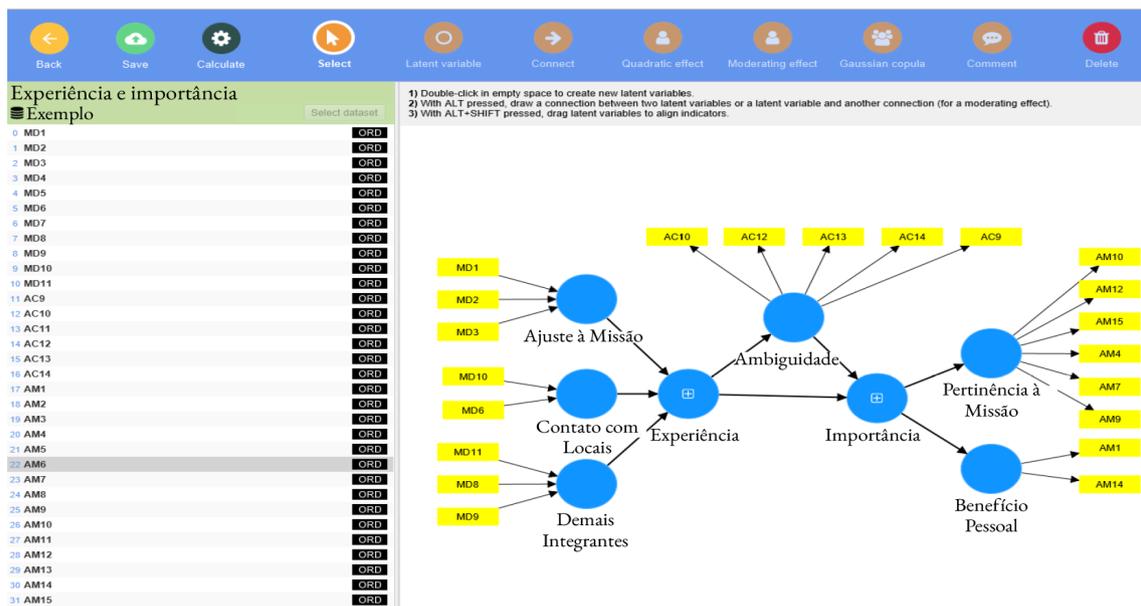
Figura 4 – Passos iniciais para a configuração dos modelos de medida e estrutural



Fonte: Elaborado pelos autores, 2022

Para configuração das variáveis latentes de segunda ordem, selecione todas as variáveis observáveis que formam/refletem as variáveis latentes de primeira ordem. Nomeie e escolha a opção *hide indicators* para que o modelo não fique poluído. Repita essa operação até que todas as variáveis latentes de seu modelo (de primeira e segunda ordem) estejam devidamente caracterizadas. Ao finalizar essa parte, clique no menu superior em *connect* e determine os caminhos de predição entre as variáveis (Figura 5). Concluídas essas fases iniciais, o modelo está pronto para ser analisado.

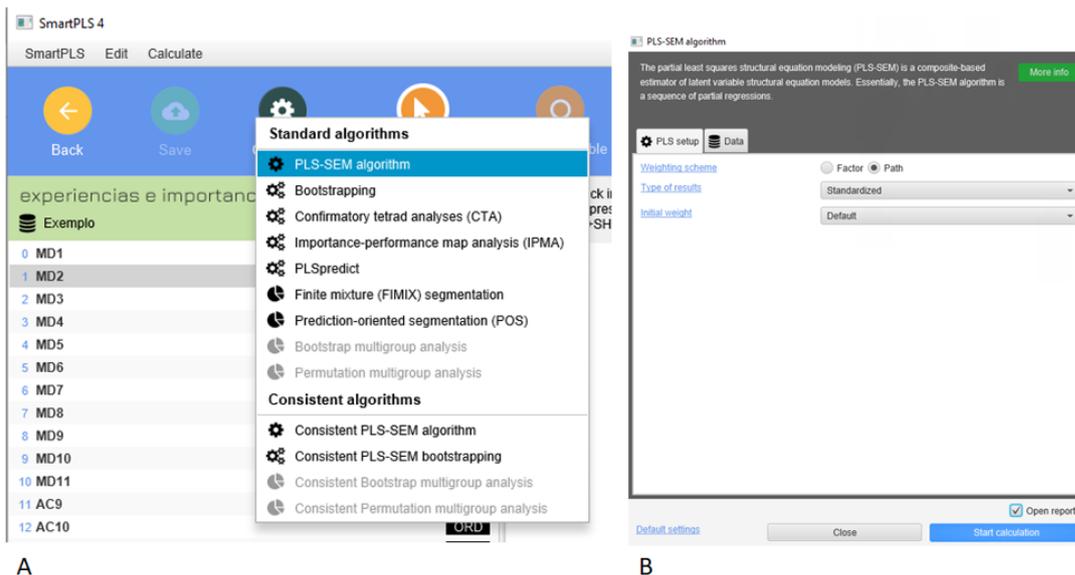
Figura 5 – Modelo pronto para análise



Fonte: Elaborado pelos autores, 2022

No exemplo em questão há modelos de medida reflexivos e formativos, a escolha foi pela primeira opção (Figura 6A). Na janela de configuração da análise, certifique-se que a opção *Path* está selecionada para *Weighting Scheme* (Figura 6B).

Figura 6 – Iniciando a análise



Fonte: Elaborado pelos autores, 2022

Uma nova janela será aberta com a saída da análise e no menu à esquerda estão disponíveis os resultados da análise. Ademais, acima da janela gráfica há opções para os quais resultados visualizar no modelo de mensuração (*outer model*) e no modelo estrutural (*inner model*) e a respeito das variáveis latentes. As informações geradas por essa análise do interesse do pesquisador estarão no tópico *quality criteria* na saída.

A análise foi iniciada pelo modelo de mensuração, conforme a recomendação da literatura. No exemplo em questão, os domínios de primeira ordem da escala de experiências (contato com locais, ajuste à missão e outros integrantes) foram modelados como formativos. Logo, nessa etapa foram inspecionados seus indicadores de variabilidade e pesos fatoriais. Todos os valores de VIF dos indicadores foram adequados. A variável observável EXP6 teve peso fatorial não significativo ($p = 0,057$), mas sua carga fatorial foi elevada ($\lambda = 0,59$). A decisão foi mantê-la no modelo, assumindo que sua importância é absoluta e não relativa (Tabela 2). Os resultados indicam que não há correlação entre indicadores (o que é altamente desejável) e que todos os indicadores presentes contribuem significativamente para formar o construto da experiência na missão de paz.

A ambiguidade de função, pertinência da missão e benefício pessoal são as variáveis latentes de primeira ordem no exemplo analisado, modeladas como construtos reflexivos. Observou-se que as variáveis observáveis AF11, da variável latente Ambiguidade de Função ($\lambda = 0,44$); IM8, da variável latente Pessoal ($\lambda = 0,46$); e IM 2 ($\lambda = 0,25$), IM3 ($\lambda = 0,11$) e IM6 ($\lambda = 0,13$) da variável latente Pertinência; tiveram cargas fatoriais baixas e foram retiradas do modelo. Esclarece-se que para retirar mais de uma variável observável de uma mesma variável latente, o pesquisador deve fazê-lo uma por uma a

cada vez, realizando uma nova estimação a cada elemento retirado. Ao retirar variáveis observáveis de uma variável latente de primeira ordem, elas devem ser retiradas da variável latente de segunda ordem. Para retirar uma variável do modelo selecione *delete* no menu superior.

Após as alterações com a remoção das variáveis observáveis indicadas, foram geradas evidências satisfatórias de validade convergente (AVE) e confiabilidade interna (confiabilidade de construto). Esses resultados permitem afirmar que as variáveis observáveis presentes em cada variável latente de primeira ordem as refletem adequadamente e que a mensuração feita por essas escalas na amostra está livre de erros aleatórios, podendo os resultados serem considerados confiáveis (Tabela 2).

Tabela 2 – Indicadores de qualidade do modelo de mensuração para as variáveis latentes e indicadores de primeira ordem

	λ/ω^*	VIF	CC	AVE
Ambiguidade de função	0,51 – 0,83	n/a	0,84	0,52
Benefício pessoal/carreira (IMP)	0,55 – 0,86	n/a	0,73	0,57
Pertinência da missão (IMP)	0,49 – 0,97	n/a	0,89	0,59
Locais (EXP)	*0,40 – 0,83	1,23 – 1,39	n/a	n/a
Ajuste à Missão (EXP)	*0,36 – 0,57	1,06 – 1,61	n/a	n/a
Outros <i>peacekeepers</i> (EXP)	*0,46 – 0,51	1,20 – 1,43	n/a	n/a

Nota: IMP = importância da missão; EXP = Experiências na missão de paz; λ = carga fatorial; ω = peso fatorial; VIF – indicador de multicolinearidade; CC = confiabilidade composta; AVE = variância média extraída; n/a = não se aplica. Valores marcados com * indicam que são referentes à ω .

Fonte: Elaborado pelos autores, 2022

Em seguida, foram inspecionadas as evidências de validade discriminante, apenas para as variáveis modeladas como reflexivas. Foram encontradas evidências satisfatórias a partir dos dados analisados (Tabela 3), ou seja, as diferentes variáveis latentes de primeira ordem inseridas no modelo são realmente distintas entre si para os dados dessa amostra.

Tabela 3 – Valores do teste HTMT para as variáveis latentes de primeira ordem

	1	2	3
1) Ambiguidade de função	-		
2) Benefício pessoal/carreira (IMP)	0,77	-	
3) Pertinência da missão (IMP)	0,72	0,30	-

Nota: IMP = importância da missão

Fonte: Elaborado pelos autores, 2022

É interessante também trazer como evidência de validade discriminante a análise das cargas cruzadas (*cross loadings*) (Figura 7), identificando os itens que carregam em cada variável latente prevista.

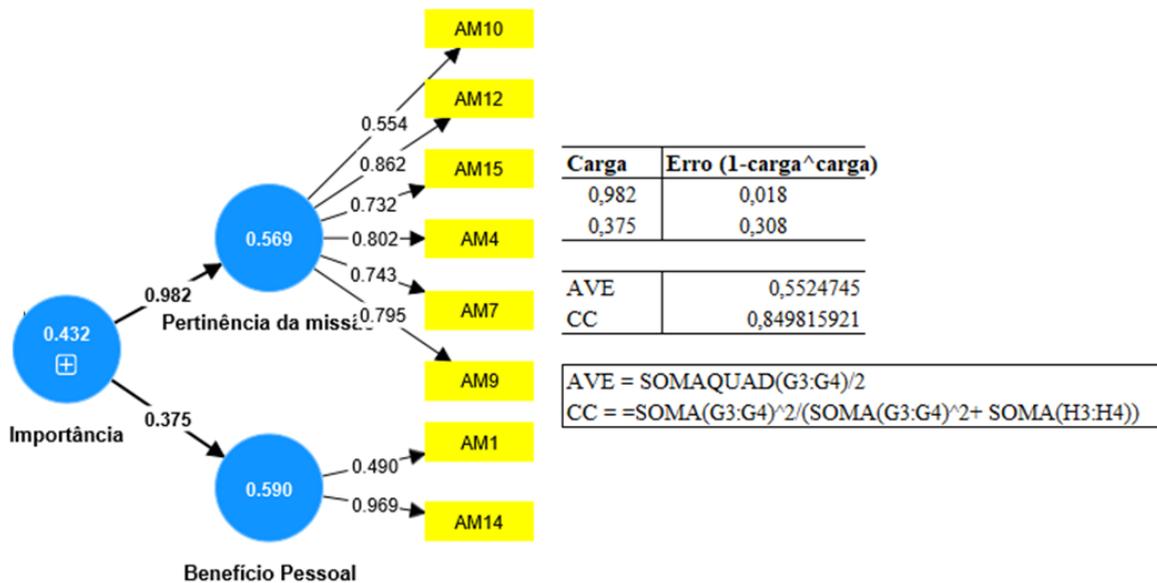
Figura 7 – Tabela com as cargas fatoriais cruzadas

	Ambiguidade	Benefício Pessoal	Pertinência da missão
AF10	0.734	-0.548	-0.328
AF12	0.734	-0.316	-0.389
AF13	0.932	-0.391	-0.472
AF14	0.768	-0.264	-0.791
AF9	0.510	-0.371	-0.109
IM1	-0.120	0.490	0.017
IM10	-0.551	0.199	0.554
IM12	-0.619	0.199	0.862
IM14	-0.512	0.969	0.213
IM15	-0.473	0.160	0.732
IM4	-0.484	0.167	0.802
IM7	-0.456	0.151	0.743
IM9	-0.401	0.026	0.795

Fonte: Elaborado pelos autores, 2022

No modelo, há uma variável latente de segunda ordem mensurada como modelo formativo: importância da missão. Há também um domínio de segunda ordem, experiências no terreno. Para o primeiro, foi calculada a AVE, a confiabilidade composta. Esses cálculos devem ser feitos à mão (ou em uma tabela Excel), pois o software fará os cálculos com as variáveis observadas repetidas, e não com as cargas das variáveis latentes de primeira ordem (Figura 8).

Figura 8 – Cálculo da AVE e Confiabilidade Composta na variável latente de segunda ordem



Fonte: Elaborado pelos autores, 2022

Não foi feito o teste HTHM, pois não há outra variável latente mensurada como modelo reflexivo para distinguir. Caso houvesse, seria necessário. Para o domínio das experiências no terreno, foi observada a multicolinearidade entre os domínios de primeira ordem (Tabela 4).

Tabela 4 – Indicadores de qualidade das variáveis latentes e domínio de segunda ordem

	VIF	CC	AVE
1) Importância da missão	n/a	0,85	0,55
2) Experiências no terreno	1,32 – 1,36	n/a	n/a

Nota: VIF – indicador de multicolinearidade; CC = confiabilidade composta; AVE = variância média extraída; n/a = não se aplica

Fonte: Elaborado pelos autores, 2022

Todos os valores mostraram-se satisfatórios, indicando que as variáveis de segunda ordem estão devidamente refletidas por variáveis de primeira ordem adequadas (validade convergente) e geraram dados confiáveis (confiabilidade interna), sendo que o construto formativo não guarda correlações indesejadas (multicolinearidade). Tendo os valores sido satisfatórios, passou-se à análise do modelo estrutural, nos qual se testa as hipóteses do estudo.

Antes de avaliar o modelo estrutural, é importante testar a colinearidade. Com esse propósito, foram analisados os valores do Fator de Inflação de Variância (VIF) para cada uma das variáveis latentes do modelo estrutural, ainda na saída da análise. Todos os valores estão dentro dos estabelecidos por Hair *et al.* (2017), abaixo de 5 (Tabela 5)

A técnica de *bootstrapping* foi utilizada para investigar a significância dos indicadores. Ela pode ser escolhida no menu superior, em *calculate*. Na aba *PLS setup* certifique-se que *path* está selecionado no *weighting scheme*. Do mais, os ajustes padrão podem ser, geralmente, mantidos. O uso do *bootstrapping* para analisar a significância das cargas adquiridas para as variáveis observáveis baseia-se não apenas em uma estimativa de modelo, mas também no cálculo de estimativas de parâmetros e seus intervalos de confiança com base em múltiplas estimativas (HAIR *et al.*, 2017). As informações de interesse estão no tópico *final results* da saída da análise de *bootstrapping*.

Tabela 5 – Efeitos diretos, indiretos específicos e totais

Efeito	Relação estrutural	VIF	β	t	p
Direto	Ambiguidade de função → Importância da missão	0,00	-0,74	8,41	<0,001
Direto	Experiência no terreno → Ambiguidade de função	1,00	0,26	1,96	0,05
Direto	Experiência no terreno → Importância da missão	0,00	-0,10	0,94	0,35
Indireto	Experiência no terreno → Ambiguidade de função → Importância da missão	n/a	0,20	1,91	0,06
Total	Experiência no terreno → Importância da missão	n/a	0,11	0,72	0,47

Fonte: Elaborado pelos autores, 2022

O teste T de *Student* fornece evidências para a análise da hipótese de que os coeficientes de correlação são iguais a zero. Se os resultados deste teste indicarem valores superiores a 1,96, a hipótese não é aceita e a correlação é significativa (HAIR *et al.*, 2017). O pesquisador não deve considerar aqui os valores atribuídos às variáveis latentes de primeira ordem. Na saída, os resultados de interesse estão no tópico *final results*. A Tabela 5 apresenta os valores dos coeficientes entre os construtos e seus respectivos testes t de *Student*.

Pode-se observar que há um efeito direto e estatisticamente significativo da ambiguidade de função na importância da missão ($\beta = -0,74$; $p < 0,001$) apenas. Isso indica que a ambiguidade de função tem um efeito preditivo negativo sobre a percepção da importância da missão.

Para avaliar o quanto as variáveis dependentes são explicadas pelo modelo, foi observado o coeficiente de determinação (R^2). Baseou-se na análise no estudo de Cohen (1988) que determina que valores iguais a 2%, 13% e 25%, são considerados pequenos, médios e grandes, respectivamente. De acordo com as análises, ambiguidade de função (predita pelas experiências no terreno) obteve $R^2 = 0,06$, considerado pequeno e a importância da missão obteve $R^2 = 0,51$, considerado grande. Logo, isso permite inferir que o modelo de predição explica 51% da importância dada à missão pelo militar (tanto no que versa acerca da percepção de seu valor para a carreira quanto no sentido da missão em si). Mas, a explicação da ambiguidade de função pelas experiências no terreno é muito pobre, cerca de 6%, o que indica ao pesquisador que outros fatores devem ser considerados em um modelo futuro para gerar uma explicação mais abrangente para a manifestação dessa variável.

Além de avaliar a magnitude dos valores de R^2 como critério de acurácia preditiva, é necessário avaliar o valor do tamanho de efeito das variáveis (isto é, sua importância explicativa no modelo) pelo indicador de *Cohen* (f^2). Com esse indicador se avalia se o construto, quando omitido, tem um impacto importante sobre os outros construtos endógenos. Valores de f^2 iguais a 0,02, 0,15 e 0,35 são considerados, respectivamente, como efeitos pequenos, médios e grandes (Cohen, 1988) da variável latente exógena (independentes). Os resultados indicaram que a relação Ambiguidade de função \rightarrow Importância da missão possui um tamanho de efeito grande ($f^2 = 0,98$); Experiência no terreno \rightarrow Importância da missão possui um tamanho de efeito pequeno ($f^2 = 0,02$), assim como Experiência no terreno \rightarrow Ambiguidade de função ($f^2 = 0,08$), o que é compreensível dado que estas duas últimas relações preditivas foram não significantes. Ambiguidade de função mostra-se como uma variável exógena de grande importância para o modelo. Os resultados de R^2 e de f^2 podem ser acessados após o pedido de análise do *PLS-algorithm* no tópico *quality criteria* da saída da análise.

O indicador de *Stone Geisser* (Q^2) é um indicador da relevância preditiva do modelo. Ele permite avaliar a contribuição de um construto exógeno para o valor Q^2 de uma variável latente endógena. Ele pode ser calculado a partir da técnica *blindfolding* ou *PLS-predict*, que se baseia em uma série de interações. O SmartPLS4 realiza apenas a última técnica (a terceira versão do *software* ainda faz as duas). Ele pode ser escolhido no menu superior, em *calculate*. Os resultados de interesse estarão no tópico *final results* da saída (*LV prediction summary*). Especificamente, quando um modelo PLS-SEM apresenta relevância preditiva, ele prediz com precisão as variáveis endógenas do modelo. Valores iguais ou superiores a zero indicam a acurácia do modelo (Tabela 6).

Tabela 6 – Valor do Indicador Stone-Geisser para as variáveis endógenas do modelo

Variáveis endógenas	MAE	Q ² _predict
Ambiguidade de função	0,82	0,04
Importância da missão	0,83	0,00

Nota: MAE = média do erro absoluto; Q² = indicador de Stone Geisser.

Fonte: Elaborado pelos autores, 2022

Neste exemplo de processo de estimação de uma PLS-SEM foram trazidas apenas as análises principais para a avaliação do modelo de medida e estrutural³. Manteve-se o banco de dados pequeno e fez-se um modelo simples para que o pesquisador iniciante nessa técnica estatística possa usá-lo na licença gratuita do *software*.

7 CONCLUSÃO

Buscou-se apresentar uma técnica estatística multivariada que pode auxiliar pesquisadores da área de ciências militares a propor novas teorias a partir de dados quantitativos. O propósito é que este ensaio teórico-metodológico instigue outros pesquisadores a considerarem essa abordagem exploratória para explicação dos fenômenos das ciências militares, segurança internacional e defesa. Uma abordagem robusta quanto às atitudes que determinam os comportamentos do homem em relação às questões de percepção de segurança, de ameaça, projeção de poder, investimento econômico, poderá trazer benefícios não apenas para a compreensão da nossa realidade local, mas para entendermos as complexidades desses fenômenos em si.

3 O banco de dados pode ser pedido ao primeiro autor do artigo por e-mail.

REFERÊNCIAS

- AKBARIYE, H.; VAZIFEDOUST, H.; SALEH ARDESTANI, A. Presenting and Validating a Model of Business Diplomacy in International Trade Based on the Grounded Theory and Structural Equation Modeling. **Future Study Management**, Tehran, v. 28, n. 111, p. 93–111, 2018.
- BARCLAY, D.; HIGGINS, C.; THOMPSON, R. The partial least squares (PLS) approach to causal modeling: personal computer adoption and use as an illustration. **Technology Studies**, [s. l.], v. 2, n. 2, p. 285–309, 1995.
- BECKER, J.; KLEIN, K.; WETZELS, M. Hierarchical latent variable models in PLS-SEM: guidelines for using reflective-formative type models. **Long Range Planning**, Amsterdam, v. 45, n. 5-6, p. 359–394, 2012.
- BIDO, D de S.; SILVA, D. da.; SOUZA, C. A. de.; GODOY, A. S. Mensuração com indicadores formativos nas pesquisas em administração de empresas: Como lidar com a multicolinearidade entre eles? **Administração: Ensino e Pesquisa**, Rio de Janeiro, v. 11, n. 2, p. 245–269, 2010.
- BOUQUET, C.; BIRKINSHAW, J. How global strategies emerge: an attention perspective. **Global Strategy Journal**, New Jersey, v. 1, n. 3–4, p. 243–262, 2011.
- BROWN, J.D. What is construct validity. **JALT**, Tokyo, v. 4 n. 2, p. 8–12, 2000.
- COHEN, J. **Statistical power analysis for the behavioral sciences**. New York: Psychology Press, 1988.
- COHEN, J. A power primer. **Psychological bulletin**, Bethesda, v. 112, n. 1, p. 155, 1992.
- COZBY, P. C.; BATES, S. C. (ed.). **Methods in behavioral research**. New York: McGraw-Hill, 2001
- CWALINA, W.; FALKOWSKI, A.; NEWMAN, B. I. Towards the development of a cross-cultural model of voter behavior: Comparative analysis of Poland and the US. **European Journal of Marketing**, Bradford, v. 44, n. 3/4, p. 351–368, 2010.
- DIAMANTOPOULOS, A.; SIGUAW, J. A. Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. **British Journal of Management**, New Jersey, v. 17, n. 4, p. 263–282, 2006.

FORNELL, C.; LARCKER, D. F. Structural Equation Models With Unobservable Variables and Measurement Error: Algebra and Statistics. **Journal of Marketing Research**, New York, v. 18, n. 3, 1981.

GARVER, M. S.; MENTZER, J. T. Logistics research methods: employing structural equation modeling to test for construct validity. **Journal of Business Logistics**, [s. l.], v. 20, n. 1, p. 33, 1999.

HAIR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. **Análise multivariada de dados**. Porto Alegre: Bookman, 2009.

HAIR, J.; RINGLE, C. M.; SARSTEDT, M. PLS-SEM: Indeed a silver bullet. **Journal of Marketing Theory and Practice**, Abingdon, v. 19, n. 2, p. 139–152, 2011.

HAIR, J. F.; SARSTEDT, M.; PIEPER, T. M.; RINGLE, C. M. The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. **Long Range Planning**, Amsterdam, v. 45, n. 5–6, p. 320-340, 2012.

HAIR, J. F.; HULT, G. T. M.; RINGLE, C. M.; SARSTEDT, M. **A primer on partial least squares structural equation modeling (PLS-SEM)**. Thousand Oaks: Sage, 2014.

HAIR, J. F.; SARSTEDT, M.; RINGLE, C. M.; GUDERGAN, S. P. **Advanced issues in partial least squares structural equation modeling**. Thousand Oaks: Sage, 2017.

HAIR, J. F.; ANDERSON, R.; BABIN, B. **Multivariate data analysis**. 18. ed. Hampshire: Cengage Learning, 2019.

HENSELER, J.; RINGLE, C. M.; SARSTEDT, M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. **Journal of the Academy of Marketing Science**, New York, v. 43, n. 1, p. 115–135, 2015.

HERSHBERGER, S.; MARCOULIDES, G.; PARRAMORE, M. Structural Equation Modeling: an introduction. **In**: PUGESEK, B. H.; TOMER, A.; VON EYE, A. (ed.). **Structural equation modeling: applications in ecological and evolutionary biology**. Cambridge: Cambridge University Press, 2003.

HURTZ, C.; BANDEIRA, D.; TRENTINI, C. (ed.). **Psicometria**. Porto Alegre: Artmed, 2015.

KARASAWA, M. Patriotism, nationalism, and internationalism among Japanese citizens: An etic–emic approach. **Political Psychology**, New York, v. 23, n. 4, p. 645–666, 2002.

KLEM, L. Structural Equation Modeling. **In:** GRIMM, L. G.; YARNOLD, P. R. **Reading and understanding MORE multivariate statistics**. Washington, DC: American Psychological Association, 2006.

KLINE, R. B. **Principles and practice of structural equation modeling**. New York: Guilford, 2015.

KOCK, N. Minimum sample size estimation in PLS-SEM: an application in tourism and hospitality research. **In:** ALI, F.; RASOOLIMANESH, S. M.; COBANOGLU, C. (ed.). **Applying partial least squares in tourism and hospitality research**. Bingley: Emerald Publishing, 2018.

KOCK, N.; HADAYA, P. Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. **Information Systems Journal**, New Jersey, v. 28, n. 1, p. 227–261, 2018.

KOUVELIS, P.; MUNSON, C. L. Using a structural equations modeling approach to design and monitor strategic international facility networks. **In:** SIMCHI-LEVI, D.; WU, S. D.; SHEN, Z-J. **Handbook of Quantitative Supply Chain Analysis**. New York: Springer, 2004.

LATAN, H.; RAMLI, N. A. The results of partial least squares-structural equation modelling analyses (PLS-SEM). **SSRN**, Amsterdam, 2013. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2364191. Acesso em: 9 fev 2022.

LOHMÖLLER, J. B. **Latent Variable Path Modeling with Partial Least Squares**. Heidelberg: Physica-Verlag, 1989.

MARCOULIDES, G. A.; SAUNDERS, C. Editor's comments: PLS: a silver bullet? **MIS Quarterly**, New York, v. 30, n. 2 p. 3–9, 2006.

MARÔCO, J. **Análise de equações estruturais: Fundamentos teóricos, software & aplicações**. Lisboa: Report Number, 2021.

MCINTOSH, C. N.; EDWARDS, J. R.; ANTONAKIS, J. Reflections on partial least squares path modeling. **Organizational Research Methods**, Thousand Oaks. v. 17, n. 2, p. 210–251, 2014.

NEVES, A. N. **Ser da paz não é fácil: análise das atitudes dos militares brasileiros em relação às missões de paz a partir de uma perspectiva microdinâmica**. 2022. Dissertação (Mestrado em Segurança Internacional e Defesa) – Escola Superior de Guerra, Rio de Janeiro, 2022.

NUNNALLY, J.; BERNSTEIN, I. **Psychometric Theory**. New York: McGraw-Hill, 1994.

PASQUALI, L. Validade dos testes psicológicos: será possível reencontrar o caminho? **Psicologia: Teoria e Pesquisa**, São Paulo, v. 23, n. spe, p. 99–107, 2007.

PETERS, G. The alpha and the omega of scale reliability and validity: why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. **European Health Psychologist**, [s. l.], v. 16, n. 2, p. 56–69, 2014.

RAHIMI SHEIKH, H.; SHARIFI, M.; SHAHRIARI, M. R. Designing a Resiliense Supply Chain Model (Case Study: The Welfare Organization of Iran). **Journal of Industrial Management Perspective**, Tehran, v. 7, n. 3, p. 127–150, 2017.

RINGLE, C. M.; DA SILVA, D.; BIDO, D de S. Modelagem de equações estruturais com utilização do SmartPLS. **Revista Brasileira de Marketing**, São Paulo, v. 13, n. 2, p. 56–73, 2014.

ROBERTS, N.; THATCHER, J. Conceptualizing and testing formative constructs: Tutorial and annotated example. **The Data Base for Advances in Information Systems**, [s. l.], v. 40, n. 3, p. 9–39, 2009.

RÖNKKÖ, M. The effects of chance correlations on partial least squares path modeling. **Organizational Research Methods**, Thousand Oaks, v. 17, n. 2, p. 164–181, 2014.

RÖNKKÖ, M.; EVERMANN, J. A critical examination of common beliefs about partial least squares path modeling. **Organizational Research Methods**, Thousand Oaks, v. 16, n. 3, p. 425–448, 2013.

RUDOLPH, A.; SCHNEIDER, F. International human trafficking: Measuring clandestinity by the structural equation approach. **IZA Discussion Papers**, Bonn, n. 7867, p. 1–47, 2013.

SARSTEDT, M.; HAIR, J. F.; RINGLE, C. M.; THIELE, K. O.; GUDERGAN, S. P. Estimation issues with PLS and CBSEM: where the bias lies! **Journal of Business Research**, Amsterdam, v. 69, n. 10, p. 3998–4010, 2016.

SCHUMACKER, R. E.; LOMAX, R. G. **A beginner's guide to structural equation modeling**. New Jersey: Lawrence Erlbaum, 2004.

SIJTSMAN, K. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. **Psychometrika**, New York, v. 74, n. 1, p. 107, 2009.

TAVAKOL, M.; DENNICK, R. Making sense of Cronbach's alpha. **International Journal of Medical Education**, Bethesda, v. 2, p. 53, 2011.

TEO, T. S. H.; SRIVASTAVA, S. C.; JIANG, L. I. Trust and electronic government success: An empirical study. **Journal of Management Information Systems**, Oxfordshire, v. 25, n. 3, p. 99-132, 2008.

URBINA, S. **Essentials of Psychological Testing**. Hoboken: John Wiley & Sons, 2004.

VINZI, V.; CHIN, W. W.; HENSELER, J.; WANG, H. **Handbook of partial least squares**. New York: Springer, 2010.