

## ELABORAÇÃO E VALIDAÇÃO DE QUESTIONÁRIOS QUANTITATIVOS: UM MODELO PRESCRITIVO

## DEVELOPMENT AND VALIDATION OF QUANTITATIVE QUESTIONNAIRES: A PRESCRIPTIVE MODEL

Paula Maria Rattis Teixeira\*  E-mail: [paula.teixeira@ufla.br](mailto:paula.teixeira@ufla.br)

\*Universidade Federal de Lavras (UFLA), São Sebastião do Paraíso, MG, Brasil.

**Resumo:** De maneira recorrente, pesquisadores do campo da Engenharia de Produção desenvolvem pesquisas que envolvem coleta de dados junto a indivíduos no âmbito organizacional. Para tanto, lançam mão de instrumentos de pesquisa que nem sempre passaram por um processo confiável de construção baseado em procedimentos fundamentados de elaboração e validação. Este artigo tem por objetivo apresentar um modelo prático, na forma de um guia prescritivo, que auxilie pesquisadores deste campo do conhecimento (e de outros correlatos) a elaborar questionários quantitativos em escala Likert focados em capturar percepções, atitudes e comportamentos de trabalhadores no contexto das organizações. Tal modelo foi elaborado a partir de revisão da literatura sobre a temática e compreende dois processos: a elaboração e posterior validação do instrumento de pesquisa. Propõem-se o emprego de três tipos de validação: de conteúdo, de fidedignidade e, por fim, a validação baseada na estrutura interna por meio da técnica estatística de análise dimensional. Ao final do artigo, um tutorial explicativo sobre como realizar uma análise dimensional no software estatístico IBM SPSS® é apresentado. Com este artigo espera-se que pesquisadores do campo da Engenharia de Produção acessem um referencial simplificado, mas que lhes permitam a construção de instrumentos de coleta de dados adequados em termos metodológicos.

**Palavras-chave:** Metodologia de pesquisa. Pesquisa quantitativa. Análise fatorial. Coleta de dados. Engenharia de Produção.

**Abstract:** In the field of Production Engineering, researchers frequently conduct studies that involve data collection from individuals within organizational settings. To this end, they often use research instruments that have not always undergone a rigorous construction process based on established development and validation procedures. This article aims to present a practical model, in the form of a prescriptive guide, to assist researchers in this field (and related areas) in designing Likert-scale-based quantitative questionnaires to capture workers' perceptions, attitudes, and behaviors in organizational contexts. The proposed model was developed based on a literature review on the subject and comprises two processes: the development and subsequent validation of the research instrument. The use of three types of validation is proposed: content validation, reliability validation, and, finally, internal structure validation through the statistical technique of dimensional analysis. At the end of the article, an explanatory tutorial on how to perform dimensional analysis using the IBM SPSS® statistical software is provided. This article seeks to offer researchers in the field of Production Engineering a simplified yet methodologically framework for constructing data collection instruments.

**Keywords:** Research methodology. Quantitative research. Factor analysis. Data collection. Production Engineering.

## 1 INTRODUÇÃO

A Engenharia de Produção é um campo que estuda sistemas produtivos desde sua implantação e operação buscando constantemente oportunidades de melhorias para estes sistemas. Ao conceber, colocar em operação e manter um sistema produtivo é essencial compreender os aspectos humanos e sociais que interferem na qualidade dos resultados esperados. Justamente por este motivo, a Engenharia de Produção historicamente se debruça sobre aspectos relacionados ao trabalho, suas condições de organização e as relações que se estabelecem no plano da sociabilidade dos trabalhadores e das organizações (Teixeira, 2024). Neste sentido, é essencial ao engenheiro compreender como aspectos humanos interferem nos processos produtivos e, para tal, lançar mão de um instrumental de coleta de dados apropriado. Não é incomum que pesquisadores da Engenharia de Produção e áreas correlatas se interessem por entender ou avaliar como variáveis como satisfação, motivação, engajamento, comprometimento, envolvimento, entre outras, interferem nos resultados operacionais. Tais variáveis são também chamadas de construtos, ou seja, são conceitos teóricos não observáveis diretamente (Uher, 2023). Para analisá-los é necessário, portanto, recorrer a um instrumento de pesquisa adequado capaz de realizar uma mensuração que de fato os avalie corretamente.

Este artigo busca propor um guia prático que auxilie profissionais e estudantes deste campo e de outros campos correlatos na construção de questionários de pesquisa quantitativos que busquem capturar percepções, atitudes e comportamentos de trabalhadores no contexto das organizações. É dada ênfase aos procedimentos de validação de tais instrumentos de modo a auxiliar os pesquisadores na elaboração de instrumentos confiáveis de coleta de dados que, de fato, mensurem os construtos que se propõem a avaliar. Para tanto, recorre-se às disciplinas associadas à Psicologia Social, mais especificamente à Avaliação Psicológica, que é uma área da Psicologia que produz hipóteses ou diagnósticos, tanto sobre pessoas quanto acerca de grupos (Hutz, 2015; Noronha *et al.*, 2023).

## 2 QUESTIONÁRIOS DE PESQUISA FOCADOS EM COMPORTAMENTOS E ATITUDES

A teoria dos testes em Psicologia se baseia no pressuposto de que um teste, um questionário ou um instrumento de medição não avalia um objeto, mas propriedades particulares de um objeto; não se avalia, por exemplo, o ser humano em sua completude, mas alguns aspectos específicos da natureza humana (Hutz, 2015; Noronha *et al.*, 2023). Em Psicologia Social e das Organizações, os instrumentos de avaliação mais utilizados são as escalas psicométricas do tipo *Likert*. A principal característica das escalas psicométricas reside no fato de que os itens de uma escala/teste refletem comportamentos ou atitudes e não o construto em si (Hutz, 2015; Pacico, 2015; Pasquali, 1998). Pasquali (1998), de forma didática, utiliza como exemplo uma avaliação sobre o construto “caridade”. O autor explica que para se avaliar a presença ou não dessa característica em um determinado indivíduo não se deve perguntar diretamente ao participante se este é caridoso, mas fazer uma série de questionamentos a respeito de comportamentos e atitudes que permitirão ao pesquisador tirar conclusões acerca da existência dessa característica no indivíduo. Essa característica/construto que não pode ser diretamente mensurada, mas depende da avaliação de comportamentos, é também chamada de traço latente. Por se basear na avaliação de comportamentos, muitas vezes, as escalas psicométricas são também chamadas de escalas de atitude ou escalas de comportamento.

Pasquali (1998) aponta que um dos principais problemas com a construção das escalas se concentra quando o pesquisador elabora uma nova escala baseado fundamentalmente em sua intuição em detrimento da literatura acerca do construto, ou seja, falta-lhe “*face validity*” ou validade aparente. Um segundo problema, ainda segundo o autor, é a falta de nitidez ou completude acerca do que exatamente a escala está medindo, ou seja, imagina-se estar medindo um construto complexo como um todo, quando, na verdade, avalia-se apenas uma parcela deste. Dessa forma, um momento essencial e, por vezes negligenciado, no processo de elaboração de um instrumento de coleta de dados é a consulta à literatura prévia acerca do construto que se pretende avaliar. As orientações aqui desenvolvidas para a construção de questionários de pesquisas se baseiam principalmente nas orientações para elaboração de escalas psicométricas.

Duas características fundamentais envolvidas no processo de construção de questionários são a validade e a fidedignidade. O conceito de validade qualifica o quanto o instrumento é legítimo para avaliar aquilo que se propõe a medir (Pacico *et al.*, 2015). Já a fidedignidade – também chamada de consistência interna – refere-se à capacidade do questionário em produzir os mesmos resultados em diferentes momentos do tempo, mas também diz respeito à habilidade do instrumento em distinguir participantes em termos de diferenças individuais (Doval; Viladrich; Angulo-Brunet, 2023; Zanon; Hauck Filho, 2015). Pasquali (1998) aponta que o pesquisador, ao elaborar um novo questionário, se vê diante de um dilema que opõe fidedignidade e funcionalidade. A fidedignidade depende diretamente do número de itens/questões do questionário, ou seja, quanto maior o número de itens, maior a garantia de que o instrumento possui capacidade de distinguir os participantes adequadamente. No entanto, longos questionários podem ser cansativos e pouco funcionais em termos de aplicabilidade.

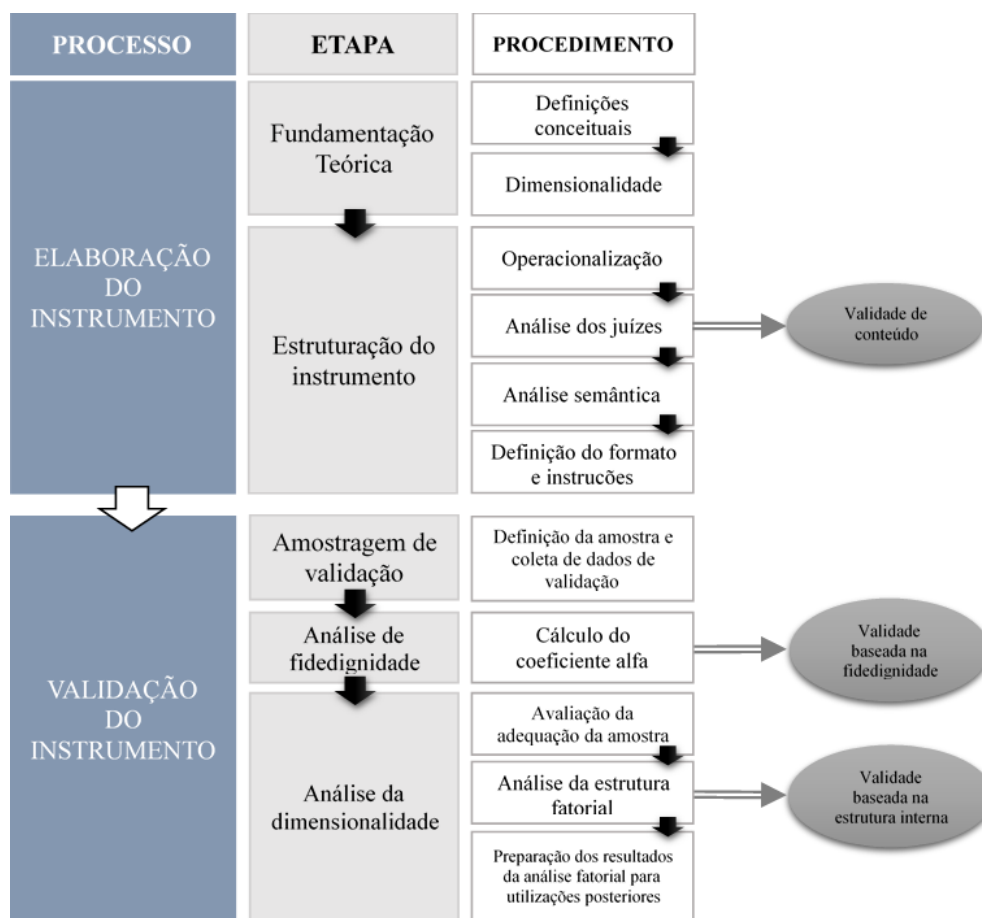
As questões elencadas anteriormente são alguns dos elementos que o pesquisador ou pesquisadora deve ter em mente no momento de escolher entre construir seu próprio instrumento de coleta de dados ou optar por algum já disponível na literatura. Uma outra questão primordial reside na teoria que fundamenta o instrumento de coleta de dados. O pesquisador deve estar ciente de que, ao escolher um instrumento já pronto, estará optando antes de tudo, por um referencial teórico específico.

### **3 MODELO PARA CONSTRUÇÃO DE QUESTIONÁRIOS DE PESQUISA**

A construção de questionários envolve basicamente dois processos: a elaboração do instrumento e sua posterior validação. Na área da Psicologia, alguns órgãos regulamentam os processos de construção e adaptação de escalas. Entre eles estão a Resolução nº 002/2003 do Conselho Federal de Psicologia (CFP) e as orientações da *American Psychological Association* (APA), *American Educational Research Association* (AERA) e *National Council on Measurement in Education* (NCME) (Pacico, 2015). Baseados nas diretrizes desses órgãos regulamentadores, alguns autores propõem uma série de guias prescritivos acerca do processo de construção de escalas. Fundamentando-se principalmente nas diretrizes de Pasquali

(1999, 2016) e Pacico (2015), é apresentado, neste artigo, um modelo de construção de questionários, conforme Figura 1. Tal modelo engloba 2 processos que se dividem em 5 etapas, as quais, por sua vez se subdividem em 11 procedimentos. O modelo aqui proposto conta com três tipos distintos de validação: validação de conteúdo por juízes especialistas; validação baseada na fidedignidade e validação baseada na estrutura interna.

**Figura 1** – Modelo prescritivo para elaboração e validação de questionários quantitativos



Fonte: Elaboração própria a partir da adaptação de Pasquali (1999, 2016) e Pacico (2015).

### 3.1 Processo 1 – Elaboração do instrumento

#### 3.1.1 1ª Etapa – Fundamentação teórica

A primeira etapa para elaboração de um instrumento de coleta de dados consiste em uma ampla revisão da literatura acerca do construto de interesse. A

revisão da literatura deve incluir, além de livros e artigos, outros instrumentos já utilizados para avaliar o construto (Pacico, 2015). Dois procedimentos podem ser destacados nessa fase: a definição teórica do construto de interesse e a indicação da dimensionalidade.

Procedimento 1 – Estabelecer a definição conceitual do construto de interesse

Primordial, uma vez estudado o construto, é a definição conceitual do mesmo. A definição constitutiva consiste em designar o construto semanticamente a partir de outros construtos, ou seja, os construtos, que “são realidades abstratas são definidos em termos de realidades abstratas” (Pasquali, 1999, p. 45). Definir o construto em termos semânticos não se trata de um exercício trivial, mas de uma tarefa que estabelecerá os limites conceituais do construto estudado e, portanto, as fronteiras do próprio instrumento de medição.

Procedimento 2 – Determinar a dimensionalidade do construto de interesse

A dimensionalidade do construto diz respeito à sua estrutura interna, ou seja, ao número de dimensões – também chamadas de atributos – que compõem o construto. Tal definição provém da literatura, de pesquisas anteriores acerca do tema estudado e de “dados empíricos disponíveis sobre ele, sobretudo dados de pesquisas que utilizaram a análise fatorial na análise dos dados, pois é fundamental decidir se o construto é uni ou multifatorial” (Pasquali, 2016, p. 38). O pesquisador costuma se deparar com diferentes vertentes teóricas relacionadas ao construto de interesse e cada uma dessas vertentes pode apontar diferentes constituições dimensionais para este. A fim de elucidar esta questão, tomemos como exemplo o construto engajamento no trabalho que é comumente apontado como possuindo múltiplas dimensões. A vertente teórica mais difundida sobre essa temática aponta que o engajamento é composto pelas dimensões vigor, dedicação e absorção (Mazzetti *et al.*, 2023; Schaufeli *et al.*, 2002). No entanto, ao se pesquisar mais a fundo esse construto, o pesquisador poderá descobrir que existem autores que compreendem o engajamento de maneira distinta e consideram que as dimensões que o compõem são engajamento físico, engajamento emocional e engajamento cognitivo (May; Gilson; Harter, 2004); ou energia, atenção e absorção (Lupano Perugini, 2024); ou ainda emoção e ativação, absorção e performance nas tarefas (Ababneh, 2025). Pasquali (1999) argumenta que essa multiplicidade de entendimentos acerca de uma

mesma temática pode levar o pesquisador a “elaborar uma mini-teoria sobre o que ele entende pelo construto que pretende medir (...). Os dados empíricos que serão coletados através do instrumento assim construído irão decidir se sua mini-teoria tem ou não alguma consistência” (Pasquali, 1999, p. 44).

### **3.1.2 2ª Etapa – Estruturação do instrumento**

A estruturação do instrumento envolve os procedimentos de produção do questionário preliminar através da operacionalização do construto, análise dos itens por juízes, avaliação da versão final em grupos focais e deliberação acerca do formato e instruções do teste.

#### **Procedimento 3 – Operacionalização do questionário de pesquisa**

Após a revisão da literatura, o pesquisador deve traduzir o construto em comportamentos, o que Pasquali (1999) chama de operacionalização, a qual, em termos práticos, consiste no procedimento de formulação dos itens/questões, ou seja, dos comportamentos que expressam a presença e magnitude do construto. A proposta desses itens pode partir de itens de outros instrumentos, revisão da literatura, entrevistas piloto com os possíveis sujeitos da pesquisa, consulta a pesquisadores, professores, etc.

É consenso entre pesquisadores que um construto para ser bem avaliado deve ter, em média, 20 itens. Esse número, no entanto, dependerá do nível de complexidade e dimensionalidade do construto (Pacico, 2015; Pasquali, 1999). Em geral, uma dimensão, para ser minimamente representada precisa de, pelo menos, três itens (Pacico, 2015). Pacico (2015) sugere que a versão preliminar do instrumento tenha de três a cinco vezes o número de itens que se pretende chegar na versão final. Pasquali (1999), por sua vez, salienta que partir de uma rigidez tal que determine um número elevado de itens na versão preliminar do instrumento faz parte de uma perspectiva positivista de construção de instrumentos de pesquisa. Partir dessa perspectiva, na visão do autor, pode levar o pesquisador a simplesmente incluir itens arbitrariamente em seu instrumento preliminar de modo a chegar em uma quantidade considerada consensualmente razoável. Para Pasquali (1999), quando se parte de

uma construção teórica robusta, o número de itens do instrumento preliminar e do instrumento final não varia tanto.

Pasquali (1999, 2016) cita doze critérios para a construção dos itens, a saber:

1. Critério comportamental: os itens devem ser expressos em termos de comportamento e não do próprio construto ou de uma abstração.
2. Critério de desejabilidade: os itens devem expressar preferência. O participante deve ser convidado a dizer se concorda, discorda ou se considera se um dado comportamento lhe convém ou não. Não existe, nesse caso, respostas certas ou erradas, mas diferentes preferências, gostos ou modos de ser.
3. Critério de simplicidade: cada item deve expressar apenas uma ideia. Deve-se evitar fornecer razões ou justificativas para uma determinada resposta.
4. Critério de clareza: o item deve ser inteligível para pessoas com menor escolaridade que constituem a população-alvo. Deve-se dar preferência para expressões simples e evitar-se frases longas e cansativas. Frases negativas tendem a ser mais confusas que frases positivas: “é melhor afirmar a negatividade do que negar uma afirmação” (Pasquali, 2016, p. 43).
5. Critério da relevância: o item deve estar alinhado à dimensão que mede. A carga fatorial encontrada na análise estatística posterior ajuda a fornecer uma compreensão acerca da relevância e pertinência do item à dimensão.
6. Critério da precisão: “o item deve ter uma posição definida no contínuo do atributo e ser distinto dos demais itens que cobrem o mesmo contínuo” (Pasquali, 2016, p. 44).
7. Critério da variedade: utilizar diferentes verbos e expressões para não cansar o respondente e utilizar itens com conotação positiva e negativa conjuntamente para evitar respostas estereotipadas.
8. Critério da modalidade: evitar expressões extremas uma vez que a intensidade do comportamento é medida na escala.
9. Critério da tipicidade: utilizar expressões ou adjetivos que se alinhem ao comportamento.

10. Critério da credibilidade: deve-se evitar que o item pareça infantil ou mesmo ridículo, de modo a não irritar ou mesmo criar má vontade do participante no momento de responder ao questionário.
11. Critério da amplitude: o conjunto de itens que formarão um determinado atributo (dimensão) “deve cobrir toda a extensão de magnitude do contínuo desse atributo” (Pasquali, 2016, p. 45).
12. Critério do equilíbrio: para um mesmo atributo é interessante conceber itens que representem comportamentos fracos, moderados e extremos, cobrindo todos os seguimentos do continuum do atributo.

A fim de exemplificar o procedimento de operacionalização/elaboração dos itens, tomemos como exemplo um questionário concebido para avaliar a presença de *burnout* entre trabalhadores, o *Burnout Assessment Tool* (BAT). Este questionário foi elaborado e validado e sua versão final conta com cinco dimensões, entre as quais uma delas é chamada de “declínio do controle cognitivo” (De Beer *et al.*, 2024; Hadžibajramović; Schaufeli; De Witte, 2022; Redelinghuys; Morgan, 2023). Na versão completa deste instrumento, esta dimensão conta com 5 itens de avaliação em que o participante deve responder com que frequência se sente de um determinado modo no seu trabalho: nunca, raramente, algumas vezes, com frequência ou sempre. Os itens são: “Em meu trabalho, tenho dificuldade em manter o foco”; “No trabalho, eu me esforço para pensar claramente”; “Sou esquecido e distraído no trabalho”; “Tenho dificuldade em me concentrar quando estou trabalhando”; e “Cometo erros no trabalho porque minha mente está em outras coisas” (Sinval *et al.*, 2022). Ao se avaliar estes itens é possível perceber que essa dimensão está expressa de modo alinhado aos critérios elencados por Pasquali (1999, 2016): foram elaborados de maneira a expressar um comportamento; não existem respostas certas ou erradas, mas o participante deve respondê-los em uma escala de intensidade; os itens expressam uma única ideia sem buscar justificá-las, são compreensíveis, relevantes e precisos; são utilizados diferentes verbos e adjetivos; expressões extremas são evitadas e é possível perceber que existe uma amplitude com itens que representam comportamentos leves a mais intensos de declínio do controle cognitivo.

Procedimento 4 – Realizar a validação de conteúdo a partir da análise dos itens por juízes especialistas

Quando o instrumento estiver pronto ele deve ser encaminhado para a apreciação dos juízes. Este procedimento tem por objetivo avaliar a adequação dos comportamentos (itens) ao atributo latente, o que implica que os juízes sejam peritos na área (Pasquali, 1999). A análise e julgamento dos itens por esses especialistas garante, juntamente com uma sólida fundamentação teórica quando da elaboração dos itens, a chamada validade de conteúdo (Alexandre; Coluci, 2011). Pacico (2015) sugere pelo menos dois juízes, mas salienta que utilizar três avaliadores pode permitir um desempate. Alexandre e Coluci (2011), no entanto, após analisarem o que distintos autores propõem na etapa de validação de conteúdo, elucidam que não há um consenso sobre o número ideal de especialistas a se adotar, com sugestões que variam de três a vinte juízes, sendo que o que deve pautar essa decisão são as características do próprio instrumento, a qualificação dos juízes e a disponibilidade de profissionais avaliadores<sup>1</sup>.

Uma tabela que contenha a definição do construto, suas dimensões, e um exemplo de como realizar a avaliação deve ser encaminhada a cada um dos juízes, juntamente com os itens. Cada item deve ser classificado pelos juízes de acordo com a sua adequação e capacidade de avaliar o construto e seus atributos. Também é interessante que haja um campo de sugestões para que cada juiz deixe sua impressão acerca da compreensibilidade, pertinência ou possíveis alterações para o item (Pacico, 2015). Alexandre e Coluci (2011) sugerem, para avaliação dos itens, uma escala ordinal que os julgue entre 1 e 4, em que 1 representa a irrelevância ou não representatividade do item para mensuração do construto e 4 representa relevância e representatividade. Ao final desse processo é possível então calcular o Índice de Validade de Conteúdo (IVC) através da seguinte fórmula:

$$IVC = \frac{\text{Número de respostas "3" ou "4"}}{\text{Número total de respostas}}$$

Alexandre e Coluci (2011) recomendam que sejam incorporados ao instrumento aqueles itens cujo IVC seja de 0,90 ou mais. Na prática, no entanto, a

---

<sup>1</sup> Em caso de pesquisas em âmbito brasileiro, sugere-se a busca de pesquisadores qualificados a partir de uma busca por termos-chave na Plataforma Lattes, que é uma base de dados de armazenamento de currículos e dados acadêmicos de pesquisadores brasileiros e estrangeiros gerenciada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Pode-se, por exemplo, elencar como possíveis juízes de conteúdo, pesquisadores com um ou mais trabalhos publicados cujo tema coincida com a temática de interesse do questionário que esteja em elaboração.

depende do número de juízes, um IVC de 0,90 implica em unanimidade do item entre todos os avaliadores. Neste caso, neste artigo, recomenda-se como satisfatório um IVC igual ou superior a 0,80 (McKechnie; Fisher, 2022; Pereira et al., 2023). Após este procedimento, pode-se dizer que o instrumento está validado em termos de conteúdo (Alexandre; Coluci, 2011).

#### Procedimento 5 – Realizar a avaliação semântica dos itens

O próximo procedimento consiste em avaliar essa nova versão por grupos focais compostos por pessoas que representem o público-alvo final (Pacico, 2015). A intenção é que os participantes avaliem o quanto os itens são compreensíveis. Pode-se solicitar que estes verbalizem o que entendem por cada item, a fim de verificar se todos compreendem de forma similar o que está sendo perguntado. Ao verificar que um item não é suficientemente compreensível, o pesquisador deve alterá-lo ou eliminá-lo do instrumento de coleta de dados (Pacico, 2015). Mais de um grupo focal pode ser conduzido até se chegar à versão final do instrumento. Ao final das análises semântica e dos juízes, espera-se ter desenvolvido um instrumento de coleta de dados que seja compreensível e representativo do construto de interesse.

#### Procedimento 6 – Definir o formato e elaborar as instruções do questionário

Uma vez que o questionário foi elaborado e estruturado, define-se o número de pontos da escala Likert que acompanhará o instrumento. Segundo Pasquali (1999), o número de pontos varia de acordo com a escolha do pesquisador, sendo os mais comuns aqueles com 5 ou 7 pontos, devendo-se, no entanto, evitar aqueles com 2 ou 3 pontos, por oferecerem poucos graus de liberdade. Acima desses valores, o número de pontos adotado pouco afeta a consistência interna, estabilidade e validade da escala Likert (Pasquali, 1999). Também nesse momento é importante redigir as instruções que acompanharão o questionário. Estas devem ser as mais compreensíveis e sucintas possíveis. Caso julgue pertinente, o pesquisador pode incluir um exemplo sobre como responder o questionário apropriadamente.

### **3.2 Processo 2 - Validação do instrumento**

O conceito de validade diz respeito ao quanto o teste é legítimo para avaliar aquilo que se propõe a medir (Pacico *et al.*, 2015; Pasquali, 1999, 2007). Pasquali (2007) salienta que enquanto “validade parece ser um conceito bastante simples, (...)

validação já se apresenta bastante mais complexa, porque se trata de provar cientificamente a validade do instrumento” (Pasquali, 2007, p. 104). Assim, o processo de validação de instrumentos de medida representa um desafio em termos do desenvolvimento de uma pesquisa. Enquanto em algumas pesquisas das ciências sociais aplicadas esse processo se restringe à avaliação da precisão do instrumento através do cálculo do Alfa de Cronbach<sup>2</sup>, em outras – especialmente na Psicologia –, trata-se de uma etapa considerada primordial, o que se justifica dada a sensibilidade que envolve o processo de avaliação de aspectos psicológicos e a importância de se garantir que um teste psicológico de fato meça o atributo que se propõe a medir. Pasquali (2007) aponta que a questão da validade se tornou um verdadeiro desafio em estudos – principalmente na área da Psicologia –, uma vez que tudo e nada é usado como argumento de validação de questionários. Ele cita então 31 tipos de validação que servem para cancelar um determinado teste psicológico<sup>3</sup>. Alguns, na opinião do autor, sem nenhum sentido e muitos deles sinônimos uns dos outros.

No modelo proposto neste artigo, a validade dos questionários se dá a partir de três frentes: *i*) validação de conteúdo, *ii*) fidedignidade calculada por meio da consistência interna e *iii*) validação baseada na estrutura interna. A validação de conteúdo acontece quando da elaboração do instrumento por meio da avaliação deste pelos juízes. A validação baseada na estrutura interna e o cálculo da fidedignidade dependem da aplicação do instrumento elaborado a uma amostra de validação e estão detalhados a seguir.

### **3.2.1 3ª Etapa – Amostragem de validação**

---

<sup>2</sup> Ressalta-se que, muito embora a avaliação acerca da precisão ou fidedignidade de um instrumento não se trate de um processo de validação, algumas pesquisas acabam por considerar o cálculo da consistência interna por meio do coeficiente Alfa de Cronbach um parâmetro de validade de um instrumento.

<sup>3</sup> É importante aqui, fazer alusão à validade nomológica. Esse tipo de processo de validação se dá quando se faz hipóteses ou suposições sobre as relações do construto de interesse com outros construtos (avaliados por meio de instrumentos já validados). O novo instrumento é validado nomologicamente quando essas hipóteses podem ser comprovadas através de análises estatísticas. Em Psicometria, é possível notar que a maior parte dos trabalhos busca a validação através desse instrumental teórico-analítico. A rede nomológica se baseia principalmente nos escores do teste para provar hipóteses de relacionamento entre construtos e assim provar a validade de um novo instrumento. Pasquali (2007) critica esse tipo de validação. O ponto essencial de sua argumentação está na seguinte afirmação: “validade diz respeito ao instrumento e não ao uso que se faz dos seus escores” (Pasquali, 2007, p. 106).

#### Procedimento 7 – Definir a amostra piloto e coletar os dados de validação

Uma vez elaborado o instrumento de coleta de dados, este deve ser aplicado a uma amostra inicial de modo que sua qualidade ou propriedades possam ser devidamente avaliadas (Pasquali, 1999). O recomendado é que esta amostra inicial reflita a população de interesse que o construto irá avaliar, ou seja, não se aconselha aplicar em crianças, um instrumento concebido para adultos, por exemplo; ou em estudantes, um instrumento projetado para trabalhadores. Com relação ao tamanho da amostra de validação, Pasquali (1999) enfatiza que, uma vez que as escalas de comportamento têm sua estrutura dimensional geralmente avaliada por meio de análise fatorial, um número mínimo de 200 participantes deve ser respeitado, de modo a viabilizar a aplicação dessa técnica estatística. Como regra prática para definir o tamanho da amostra inicial, o autor sugere a aplicação em 5 a 10 participantes por item, respeitando o mínimo de 200 aplicações (Pasquali, 1999).

#### **3.2.2 4ª Etapa – Análise da fidedignidade do instrumento**

##### Procedimento 8 – Calcular a consistência interna com o coeficiente Alfa

A fidedignidade é um parâmetro de um instrumento também conhecido por confiabilidade, consistência interna, estabilidade ou precisão (Doval; Viladrich; Angulo-Brunet, 2023; Zanon; Hauck Filho, 2015). Refere-se à “estabilidade com que os escores dos testandos conservam-se em aplicações alternativas de um mesmo teste ou em formas equivalentes de testes distintos” (Zanon; Hauck Filho, 2015, p. 84). Dito de outro modo, um questionário fidedigno quando aplicado a um mesmo participante em dois momentos distintos de tempo retorna resultados similares, desde que mudanças significativas não tenham se dado na realidade deste participante. É essencial que um instrumento de mensuração consiga recuperar os mesmos escores considerando que o construto avaliado apresente estabilidade ao longo do tempo, mas é também fundamental que o instrumento de coleta de dados “consiga diferenciar apropriadamente testandos com diferentes níveis no traço latente de interesse” (Zanon; Hauck Filho, 2015, p. 84).

Ao se avaliar a fidedignidade de um questionário, avalia-se o grau de consistência entre múltiplas medidas de uma variável (item) (Hair Jr. *et al.*, 2019). Trata-se, portanto, de uma propriedade fundamental para determinar a validade de

um questionário. Não é, no entanto, uma condição suficiente para uso de um determinado instrumento de coleta de dados. Normalmente se começa a validação de um questionário com a análise da fidedignidade por ser mais simples e menos trabalhosa. “Se não houver evidências de fidedignidade, também não haverá de validade” (Zanon; Hauck Filho, 2015, p. 84). O parâmetro mais usado para avaliação da fidedignidade em instrumentos de mensuração com escala *Likert* é o coeficiente alfa, também chamado Alfa de Cronbach. Trata-se de uma medida de confiabilidade que varia entre 0 e 1. Hair Jr. *et al.* (2019) sugere como aceitáveis questionários que apresentem um valor mínimo de 0,70 para este parâmetro. Diversos pacotes estatísticos possuem instrumental para calcular rapidamente o coeficiente alfa.

### **3.2.3 5ª Etapa – Análise da dimensionalidade do instrumento (validação baseada na estrutura interna)**

Esta validação do instrumento consiste em verificar quantas dimensões compõem o questionário a partir dos dados coletados na amostra de validação. Se o instrumento foi concebido como sendo uni ou multidimensional, é a partir desta análise que se poderá avaliar sua real dimensionalidade. A partir da validação baseada na estrutura interna também é possível determinar eventuais itens que não se relacionam com nenhuma dimensão e conseqüentemente com o construto e que devem, portanto, ser excluídos do instrumento de coleta de dados e das análises seguintes. Trata-se aqui de uma análise exploratória uma vez que a intenção é avaliar um questionário novo, concebido pelo pesquisador e para o qual se há pouca ou nenhuma evidência sobre a estrutura prévia dos dados.

A análise dimensional pode ser realizada a partir do emprego de diferentes técnicas. Técnicas de análise e redução de dimensão, como a análise fatorial, buscam entender como um conjunto de variáveis<sup>4</sup> correlacionadas pode se agrupar ou ser reduzido em uma menor quantidade de variáveis resumo não correlacionadas, explicando a maior quantidade possível de variância total ou comum dos dados (Hair Jr. *et al.*, 2019). Essas novas variáveis resumo são também chamadas de componentes, fatores ou dimensões, de acordo com a técnica de redução/resumo

---

<sup>4</sup> Aqui, estamos falando da análise fatorial realizada sobre um questionário composto por itens avaliados em escala *Likert*. Quando se usa o termo variáveis, está-se falando justamente dos itens que compõem o questionário.

empregada<sup>5</sup>. Técnicas de análise dimensional são técnicas de interdependência, empregadas tanto para resumir quanto para reduzir variáveis. O resumo obtém dimensões que descrevem os dados com um número menor de conceitos quando comparado com as variáveis originais. Já a redução vai além e deriva um valor empírico (escore fatorial) para cada dimensão de modo a ser utilizado em análises subsequentes. Para o emprego de tais técnicas, é desejável que haja multicolinearidade, uma vez que o objetivo é encontrar inter-relações entre as variáveis em estudo (Hair Jr. *et al.*, 2019).

A Análise de Fatores Comuns (AFC) é um tipo de técnica de redução de dimensões que busca encontrar o melhor agrupamento de variáveis – também chamado de fatores – a partir da variância comum dos dados. Neste caso, geralmente há um conhecimento prévio da estrutura dos dados. Já a Análise em Componentes Principais (ACP) é um tipo particular de técnica de agrupamento e redução, também apropriada a dados quantitativos, que gera componentes (agrupamento de variáveis originais) não correlacionadas entre si com o intuito de maximizar a variância total explicada dos dados (Greenacre *et al.*, 2022; Marôco, 2021). Tanto a ACP quanto a AFC são técnicas apropriadas para dados quantitativos. Uma alternativa, ao se trabalhar com dados qualitativos, como os gerados por escalas *Likert*, consiste no emprego da análise de componentes principais não linear. No software estatístico IBM SPSS®, essa análise recebe o nome de Análise de Componentes Principais para Dados Categóricos (CATPCA) (Campos *et al.*, 2020; IBM, 2026). As saídas da CATPCA são similares às da ACP e, no caso de se trabalhar exclusivamente com variáveis quantitativas, ambas as técnicas retornarão os mesmos resultados (Linting; Van Der Kooij, 2012).

O que se percebe ao analisar artigos científicos que buscam realizar análise dimensional de questionários é que ACP e AFC são mais comumente utilizadas. Na prática, os resultados gerados a partir das três técnicas de análise dimensional não diferem substancialmente entre si. Desta forma, ao final do artigo (Anexo 1), é apresentado um tutorial sobre como realizar a análise dimensional no software IBM SPSS® empregando a AFC. Para executá-lo o pesquisador deve possuir um

---

<sup>5</sup> Geralmente utiliza-se o termo componente para o agrupamento de variáveis resultante da ACP; fator para o agrupamento da AFC e dimensão para a CATPCA. Neste trabalho, são utilizados qualquer um dos termos indistintamente.

conhecimento básico prévio sobre como trabalhar com bancos de dados neste software.

A fim de empregar as técnicas de análise dimensional alguns pressupostos estatísticos devem ser observados e os procedimentos necessários serão discutidos a seguir.

#### Procedimento 9 – Avaliar a adequação da amostra

Como a multicolinearidade é um pressuposto desejável para a aplicação de técnicas de interdependência, é importante gerar uma matriz de correlações de Pearson de modo a observar o comportamento dos dados. Se em uma primeira análise, um número substancial de correlações não for maior que 0,30 (em números absolutos), então a análise dimensional pode não ser apropriada ao questionário (Hair Jr. *et al.*, 2019). No entanto, análises mais aprofundadas podem ser realizadas para verificar pertinência da utilização dessa técnica. Sugere-se a utilização de três diferentes parâmetros para avaliar a adequação da amostra para aplicação da análise dimensional: *i*) teste de esfericidade de Bartlett, *ii*) teste Kaiser-Meyer-Olkin (KMO) e *iii*) análise das correlações parciais da matriz anti-imagem.

O teste de esfericidade de Bartlett fornece a significância estatística de que a matriz de correlações de Pearson tem correlações significantes entre, pelo menos, algumas variáveis (Hair Jr. *et al.*, 2019; Rogers, 2022). O teste KMO retorna uma medida de adequação da amostra (MSA<sup>6</sup>) que varia entre 0 e 1, e cujos valores menores que 0,50 são considerados inaceitáveis para aplicação da análise dimensional (Hair Jr. *et al.*, 2019; Rogers, 2022). Já por meio da análise da matriz anti-imagem, é possível examinar as correlações parciais. Correlação parcial é “aquela que não é explicada quando os efeitos de outras variáveis são levados em consideração” (Hair Jr. *et al.*, 2019, p. 109). Se as correlações parciais são altas, as técnicas de análise fatorial não são adequadas (exceção quando um fator tem apenas duas variáveis). Correlações parciais maiores que 0,70 são consideradas altas (Hair Jr. *et al.*, 2019).

#### Procedimento 10 – Analisar a estrutura fatorial dos dados

---

<sup>6</sup> Do inglês *Measure of Sampling Adequacy*.

Uma vez que a amostra se mostrou adequada, pode-se realizar a análise dimensional propriamente dita. Uma questão relevante ao se trabalhar com técnicas de análise dimensional consiste no número de dimensões ou de fatores a extrair. Uma vez que um fator é uma combinação linear entre variáveis, a cada fator (ou dimensão) está associado um autovalor (Field, 2024). Uma medida usual consiste em considerar apenas aqueles fatores cujos autovalores sejam maiores que 1, o que é conhecido por critério da raiz latente (Hair Jr. *et al.*, 2019; Ruengvirayudh; Brooks, 2022; Wang *et al.*, 2025). No entanto, outros aspectos podem ser observados, como a porcentagem de variância explicada pelo modelo, estudos preliminares acerca do conceito estudado ou mesmo a heterogeneidade entre os respondentes (Hair Jr. *et al.*, 2019).

A interpretação dos fatores (dimensões) gerados parte da análise das cargas fatoriais que são as correlações entre cada variável e o fator. Quanto maior a carga, maior a capacidade explicativa da variância da variável pelo fator (Hair Jr. *et al.*, 2019). Na prática, o que acontece é que uma variável pode ter cargas significantes em mais de um fator, o que é chamado de carga cruzada (Hair Jr. *et al.*, 2019). Resultados mais interpretáveis podem ser encontrados por meio da rotação dos fatores, quando as variáveis têm suas cargas maximizadas sobre um fator (Field, 2024). O objetivo é tornar as cargas altas mais explícitas e diminuir as cargas cruzadas de modo a ter maior segurança no momento de definir quais variáveis pertencem a quais fatores (Hair Jr. *et al.*, 2019). No entanto, a decisão de utilizar dimensões rotacionadas ou não depende de algumas questões importantes e do que se deseja extrair da análise, não sendo uma unanimidade entre autores.

Se a carga fatorial é correlação entre a variável e o fator, a carga ao quadrado é “a quantia de variância total da variável explicada pelo fator” (Hair Jr. *et al.*, 2019, p. 119). Nesse caso, se uma determinada variável (item) possui uma carga de 0,30 em um determinado fator, 9% da variância dessa variável é explicada por esse fator. É baseado nesses valores que o pesquisador deve optar por manter uma determinada variável em um ou outro fator, ou mesmo excluí-la do seu instrumento de coleta de dados. Hair Jr. *et al.* (2019) sugere como regra prática que variáveis que apresentem cargas menores que 0,30 em um determinado fator sejam excluídas; cargas entre 0,30 e 0,40 atendem a um nível mínimo de interpretação da estrutura; cargas entre 0,50 e

0,70 podem ser consideradas praticamente significantes e cargas acima de 0,70 são o objetivo de qualquer análise dimensional.

Após a análise da matriz fatorial, alguns problemas podem persistir, como por exemplo uma variável não ter cargas significantes, mesmo com carga significativa possuir uma comunalidade muito baixa, ou a variável persistir com cargas cruzadas. Algumas possíveis ações corretivas seriam ignorar a variável problemática, eliminá-la do questionário, tentar novos modelos de rotação, diminuir ou aumentar o número de dimensões, ou até mesmo mudar a técnica estatística de análise (Hair Jr. *et al.*, 2019). A decisão sobre qual atitude tomar, fica a cargo da sensibilidade do pesquisador. Na prática, o que geralmente acontece é que várias análises são rodadas e ao pesquisador cabe a decisão de escolher a solução que apresente uma melhor adequação aos dados.

Após a aplicação da análise fatorial, a dimensionalidade previamente concebida no momento de elaboração do instrumento de coleta de dados pode ser confirmada ou mesmo repensada. Alguns itens que compõem cada dimensão podem ser excluídos ou mesmo realocados de dimensão. Têm-se assim, a validação baseada na estrutura interna do instrumento.

Procedimento 11 – Estabelecer como os resultados serão utilizados em análises posteriores

Uma vez definida a estrutura fatorial, é possível estabelecer uma configuração dos resultados, baseada na amostra de validação, de modo a utilizar essa nova configuração em análises subsequentes. Hair Jr. *et al.* (2019) apontam que, ao se trabalhar com análise fatorial, o pesquisador pode ter como objetivo simplesmente resumir os dados, ou seja, fazer da análise dimensional um fim em si mesma, isto é, utilizar a técnica somente para estimar os fatores e determinar as contribuições de cada variável aos fatores. No entanto, o pesquisador pode ir além e utilizar os resultados encontrados para redução, criando um novo e menor conjunto de variáveis que será utilizado em análises subsequentes, sejam elas utilizando o mesmo banco de dados ou um novo banco de dados gerado a partir da aplicação do instrumento previamente validado.

Uma das formas de trabalhar com a redução de variáveis é utilizar escalas múltiplas que são novas escalas que correspondem aos fatores gerados na análise

fatorial e são compostas pelos itens que aderiram àquele determinado fator, ou seja, apresentaram cargas altas no mesmo. Um valor único por observação e por fator pode ser calculado utilizando média aritmética das respostas aos itens (Hair Jr. *et al.*, 2019). Outra opção é utilizar média ponderada, em que, para cada observação, é calculado um valor único por fator que é a média ponderada das notas atribuídas pelos participantes aos itens que compõem aquele fator e cujos pesos ponderadores são as cargas fatoriais encontradas para aquela variável naquele fator, a partir da amostra de validação do questionário (Borges; Alves-Filho; Tamayo, 2008). Segundo Hair Jr. *et al.* (2019) é essencial verificar a unidimensionalidade das novas escalas criadas, bem como realizar testes de confiabilidade das mesmas. Sugere-se duas estimativas de confiabilidade para avaliar as novas escalas múltiplas: o Alfa de Cronbach e o coeficiente Ômega de McDonald. A escolha desse segundo coeficiente se dá em razão de que alguns autores criticam o uso do Alfa de Cronbach em análises de confiabilidade uma vez que este faz suposições estatísticas que dificilmente se concretizam na prática (Béland; Falk, 2022). Por exemplo, para o cálculo de Alfa, se considera que todas as cargas fatoriais são iguais. Dito de forma simplificada, todos os itens teriam igual capacidade de explicar o construto, o que, em geral não é verdade. Assim, não faz diferença se a análise fatorial foi rodada ou não, já que o cálculo de Alfa ignora cargas fatoriais. O Ômega de McDonald apresenta-se como alternativa, uma vez que é um parâmetro que leva as cargas em consideração para o cálculo de confiabilidade. Seu ponto de corte segue o padrão de  $\alpha$ , consideradas confiáveis escalas cujo  $\omega$  é maior que 0,70 (Revelle; Condon, 2019).

#### **4 CONSIDERAÇÕES FINAIS**

O presente artigo teve como objetivo apresentar um modelo estruturado e simplificado para elaboração de questionários quantitativos de pesquisa centrados em comportamentos humanos, especialmente no contexto organizacional. Ao propor tal modelo espera-se que pesquisadores do campo da Engenharia de Produção e outras áreas correlatas tenham um referencial simplificado, mas que lhes permitam a construção de instrumentos consistentes de coleta de dados em termos metodológicos. Dessa forma, neste artigo, nos detivemos em apresentar um modelo

estruturado que baseia a validação do questionário de pesquisa em três frentes: i) a validação de conteúdo, ii) a fidedignidade calculada por meio da consistência interna e iii) a validação baseada na estrutura interna.

Muito embora não seja a intenção deste artigo elaborar instruções sobre como criar um teste psicológico ou uma escala psicométrica (que devem ser manipulados por profissionais dessa área e com formação para tal), não se perde de vista que o modelo aqui proposto tem a intenção de fornecer subsídios para a construção de questionários válidos e capazes de capturar subjetividades e especificidades de indivíduos no contexto organizacional. Dessa forma, embora o modelo aqui proposto se baseie no referencial teórico da Psicologia Social não se discute aqui processos de validação baseados em redes nomológicas tão comuns em trabalhos dessa área, nos quais são empregados, principalmente, análise fatorial confirmatória e modelagem de equações estruturais. Caso seja do interesse do pesquisador, este pode buscar um aprofundamento nestas técnicas de validação de escalas psicométricas dado o abundante referencial teórico sobre o tema.

## AGRADECIMENTOS

Este artigo é resultado da pesquisa de doutorado da autora que foi financiada parcialmente pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (Processo 88887.464745/2019-00).

## REFERÊNCIAS

ABABNEH, O. M. How can human capital promote innovative behaviour? exploring the attitudinal dynamics of employee engagement and mental involvement. **The International Journal of Entrepreneurship and Innovation**, Londres, v. 26, n. 2, p. 108–122, 2025. DOI: <https://doi.org/10.1177/14657503221123405>

ALEXANDRE, N. M. C.; COLUCI, M. Z. O. Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. **Ciência & Saúde Coletiva**, Rio de Janeiro, v. 16, n. 7, p. 3061–3068, 2011. DOI: <https://doi.org/10.1590/S1413-81232011000800006>

BÉLAND, S.; FALK, C. F. A Comparison of Modern and Popular Approaches to Calculating Reliability for Dichotomously Scored Items. **Applied Psychological**

**Measurement**, Minneapolis, v. 46, n. 4, p. 321–337, 2022. DOI: <https://doi.org/10.1177/01466216221084210>

BORGES, L. de O.; ALVES-FILHO, A.; TAMAYO, Á. Motivação e significado do trabalho. *In*: SIQUEIRA, M. M. M. (org.). **Medidas do Comportamento Organizacional: ferramentas de diagnóstico e de gestão**. Porto Alegre: Artmed, 2008. p. 215–248.

CAMPOS, C. I. D.; PITOMBO, C. S.; DELHOMME, P.; QUINTANILHA, J. A. Comparative analysis of data reduction techniques for questionnaire validation using self-reported driver behaviors. **Journal of Safety Research**, Itasca, v. 73, p. 133–142, 2020. DOI: <https://doi.org/10.1016/j.jsr.2020.02.004>

DE BEER, L. T.; SCHAUFELI, W. B.; DE WITTE, H.; HAKANEN, J. J.; KALTIAINEN, J.; GLASER, J.; SEUBERT, C.; SHIMAZU, A.; BOSAK, J.; PROCHÁZKA, J.; KAJZAR, A.; CHRISTENSEN, M. Revisiting a global burnout score with the Burnout Assessment Tool (BAT) across nine country samples. **European Journal of Psychological Assessment**, Göttingen, p. 1–6, 2024. DOI: <https://doi.org/10.1027/1015-5759/a000839>

DOVAL, E.; VILADRICH, C.; ANGULO-BRUNET, A. Coefficient Alpha: The Resistance of a Classic. **Psicothema**, Oviedo, v. 1, n. 35, p. 5–20, 2023. DOI: <https://doi.org/10.7334/psicothema2022.321>

FIELD, A. **Discovering Statistics Using IBM SPSS Statistics**. London Thousand Oaks, California New Delhi Singapore: Sage Publications Ltd, 2024. 1144 p.

GREENACRE, M.; GROENEN, P. J. F.; HASTIE, T.; IODICE D'ENZA, A.; MARKOS, A.; TUZHILINA, E. Principal component analysis. **Nature Reviews Methods Primers**, Londres, v. 2, n. 100, 2022. DOI: <https://doi.org/10.1038/s43586-022-00184-w>

HADŽIBAJRAMOVIĆ, E.; SCHAUFELI, W.; DE WITTE, H. Shortening of the Burnout Assessment Tool (BAT)—from 23 to 12 items using content and Rasch analysis. **BMC Public Health**, Londres, v. 22, n. 560, p. 1-16, 2022. DOI: <https://doi.org/10.1186/s12889-022-12946-y>

HAIR JR., J.; BLACK, W.; BABIN, B.; ANDERSON, R. **Multivariate Data Analysis**. 8. ed. Andover: Cengage Learning, 2019. 813 p.

HUTZ, C. S. O que é avaliação psicológica: métodos, técnicas e testes. *In*: HUTZ, C. S.; BANDEIRA, D. R.; TRENTINI, C. M. (org.). **Psicometria**. Porto Alegre: Artmed, 2015. p. 11–21.

IBM. **Categorical Principal Components Analysis (CATPCA). Software documentation - 31.0.0 change version**. Armonk: IBM, 2026. Disponível em:

[https://www.ibm.com/support/knowledgecenter/en/SSLVMB\\_23.0.0/spss/categories/dh\\_cpca.html](https://www.ibm.com/support/knowledgecenter/en/SSLVMB_23.0.0/spss/categories/dh_cpca.html). Acesso em: 15 maio 2022.

LINTING, M.; VAN DER KOOIJ, A. Nonlinear principal components analysis with CATPCA: a tutorial. **Journal of Personality Assessment**, London, v. 94, n. 1, p. 12–25, 2012. DOI: <https://doi.org/10.1080/00223891.2011.627965>

LUPANO PERUGINI, M. L. Rasgos Positivos de La Tríada Luminosa como Predictores de Satisfação y Performance Laboral. El Rol Mediador del Engagement. **Revista psicodebate: psicología, cultura y sociedad**, Buenos Aires, v. 24, n. 2, p. 63–76, 2024. DOI: <https://doi.org/10.18682/pd.v24i2>

MARÔCO, J. **Análise Estatística com o SPSS Statistics**. 8. ed. Pêro Pinheiro: ReportNumber, 2021. 1022 p.

MAY, D. R.; GILSON, R. L.; HARTEK, L. M. The psychological conditions of meaningfulness, safety and availability and the engagement of the human spirit at work. **Journal of Occupational and Organizational Psychology**, Leics, v. 77, n. 1, p. 11–37, 2004. DOI: <https://doi.org/10.1348/096317904322915892>

MAZZETTI, G.; ROBLEDO, E.; VIGNOLI, M.; TOPA, G.; GUGLIELMI, D.; SCHAUFELI, W. B. Work engagement: a meta-analysis using the job demands-resources model. **Psychological Reports**, Thousand Oaks, v. 126, n. 3, p. 1069–1107, 2023. DOI: <https://doi.org/10.1177/003329412111051988>

MCKECHNIE, D.; FISHER, M. J. Considerations when examining the psychometric properties of measurement instruments used in health. **Australian Journal of Advanced Nursing**, Melbourne, v. 39, n. 2, 2022. DOI: <https://doi.org/10.37464/2020.392.481>

NORONHA, A. P. P.; REPPOLD, C. T.; BANDEIRA, D. R.; SANTOS, A. A. A. D. The Development of Psychological Assessment in Brazil and Current and Future Challenges. **Psicologia: Teoria e Pesquisa**, Brasília, v. 39, n. spe, p. e39nspe10, 2023. DOI: <https://doi.org/10.1590/0102.3772e39nspe10.en>

PACICO, J. C. Como é feito um teste? Produção de itens. *In*: HUTZ, C. S.; BANDEIRA, D. R.; TRENTINI, C. M. (org.). **Psicometria**. Porto Alegre: Artmed, 2015. p. 54–69.

PACICO, J. C.; HUTZ, C. S.; SCHNEIDER, A.; BANDEIRA, D. R. Validade. *In*: HUTZ, C. S.; BANDEIRA, D. R.; TRENTINI, C. M. (org.). **Psicometria**. Porto Alegre: Artmed, 2015. p. 70–83.

PASQUALI, L. Instrumentação no estudo das organizações: a utilização de escalas psicométricas. *In*: TAMAYO, Á.; BORGES-ANDRADE, J. E.; CODO, W. (org.).

**Trabalho, organizações e cultura.** São Paulo: Cooperativa de Autores Associados, 1998. p. 75–82.

PASQUALI, L. Princípios de elaboração de escalas. *In*: GORENSTEIN, C.; WANG, Y.-P.; HUNGERBÜHLER, I. (org.). **Instrumentos de avaliação em saúde mental.** Porto Alegre: Artmed, 2016.

PASQUALI, L. Testes referentes a construto: teoria e modelo de construção. *In*: PASQUALI, L. (org.). **Instrumentos psicológicos: manual prático de elaboração.** Brasília: LabPAM/IBAPP, 1999. p. 37–72.

PASQUALI, L. Validade dos testes psicológicos: será possível reencontrar o caminho? **Psicologia: Teoria e Pesquisa**, Brasília, v. 23, n. número especial, p. 99–107, 2007. DOI: <https://doi.org/10.1590/S0102-37722007000500019>

PEREIRA, D. S.; FLORÊNCIO, R. S.; CESTARI, V. R. F.; MOREIRA, T. M. M. Vulnerability to physical inactivity: evidence of content validity and response processes. **Revista Brasileira de Enfermagem**, Brasília, v. 76, p. e20220563, 2023. DOI: <https://doi.org/https://doi.org/10.1590/0034-7167-2022-0563>

REDELINGHUYS, K.; MORGAN, B. Psychometric properties of the Burnout Assessment Tool across four countries. **BMC Public Health**, Londres, v. 23, n. 824, p. 1-14, 2023. DOI: <https://doi.org/10.1186/s12889-023-15604-z>

REVELLE, W.; CONDON, D. M. Reliability from  $\alpha$  to  $\omega$ : A tutorial. **Psychological Assessment**, Gottingen, v. 31, n. 12, p. 1395–1411, dez. 2019. DOI: <https://doi.org/10.1037/pas0000754>

ROGERS, P. Melhores Práticas para sua Análise Fatorial Exploratória: Tutorial no Factor. **Revista de Administração Contemporânea**, Rio de Janeiro, v. 26, p. e210085, 2022. DOI: <https://doi.org/10.1590/1982-7849rac2022210085.por>

RUENGVIYAYUDH, P.; BROOKS, G. P. Sample Size for Parallel Analysis and Not-So-Common Criteria for Dimensions in Factor Analysis: Modifying the Eigenvalue > 1 Kaiser Rule. **General Linear Model Journal**, DeKalb, v. 46, n. 1, p. 20–34, 2022. DOI: <https://doi.org/10.31523/glmj.046001.002>

SCHAUFELI, W.; SALANOVA, M.; GONZALEZ-ROMA, V.; BAKKER, A. The measurement of engagement and burnout: a two sample confirmatory factor analytic approach. **Journal of Happiness Studies**, Dordrecht, v. 3, n. 1, p. 71–92, 2002.

SINVAL, J.; VAZQUEZ, A. C. S.; HUTZ, C. S.; SCHAUFELI, W. B.; SILVA, S. Burnout Assessment Tool (BAT): validity evidence from Brazil and Portugal. **International Journal of Environmental Research and Public Health**, Basileia, v. 19, n. 3, p. 1–25, 25, 2022. DOI: <https://doi.org/10.3390/ijerph19031344>

TEIXEIRA, P. M. R. **Dimensões do significado do trabalho e a constituição de perfis de trabalhadoras e trabalhadores brasileiros**: relações com aspectos sociais, demográficos e de empregabilidade. 2024. 290 f. Tese (Doutorado em Engenharia de Produção) – Universidade de São Paulo, São Carlos, 2024. DOI: <https://doi.org/10.11606/T.18.2024.tde-06022025-105327>

UHER, J. What Are Constructs? Ontological Nature, Epistemological Challenges, Theoretical Foundations and Key Sources of Misunderstandings and Confusions. **Psychological Inquiry**, Abingdon, v. 34, n. 4, p. 280–290, 2023. DOI: <https://doi.org/10.1080/1047840X.2023.2274384>

WANG, C.; GUO, J.; MA, Y.; ZHENG, S. Kaiser Criterion in Factor Models. **Acta Mathematica Sinica, English Series**, Beijing, v. 41, n. 2, p. 547–552, 2025. DOI: <https://doi.org/10.1007/s10114-025-3383-3>

ZANON, C.; HAUCK FILHO, N. Fidedignidade. *In*: HUTZ, C. S.; BANDEIRA, D. R.; TRENTINI, C. M. (org.). **Psicometria**. Porto Alegre: Artmed, 2015. p. 84–95.

### **Biografia do(s) autor(es)**

#### **Paula Maria Rattis Teixeira**

Possui graduação em Engenharia de Produção Mecânica pela Universidade de São Paulo - USP (2006), mestrado em Ciência, Tecnologia e Sociedade pela Universidade Federal de São Carlos - UFSCar (2011) e doutorado em Engenharia de Produção pela Universidade de São Paulo - USP (2024). Atualmente é Professora Adjunta da Universidade Federal de Lavras (campus de São Sebastião do Paraíso), lotada no Instituto de Ciência, Tecnologia e Inovação (ICTIN/UFLA) e Professora Visitante no Programa de Pós-Graduação em Planejamento e Análise de Políticas Públicas (PAPP/UNESP). Desenvolve pesquisas tanto quantitativas, quanto qualitativas. Suas áreas de interesse estão na interface entre organizações, processos produtivos e trabalhadores dando ênfase a pesquisas sobre as relações de trabalho e seu impacto sobre os trabalhadores no contexto de precarização das relações laborais e crise do capital.



Artigo recebido em: 01/12/2024 e aceito para publicação em: 24/05/2026  
DOI: <https://doi.org/10.14488/1676-1901.v26i2.5483>

## ANEXOS

### Anexo 1 – Guia prático para executar a análise de fidedignidade e dimensionalidade de questionário quantitativo no software estatístico IBM SPSS

	CAMINHO	INTERPRETAÇÃO DA SAÍDA	VALORES DE REFERÊNCIA
Cálculo do coeficiente alfa	<p>Analisar → Escala → Análise de confiabilidade → Escolha as variáveis → Em “Modelo” escolha a opção “Alfa” → OK</p> <p><i>OBS: Na caixa “Estatística” ao marcar a opção “escalar se o item for excluído”, a saída apresenta o valor de Alfa caso cada variável seja excluída</i></p>	<p>Verificar a saída “<b>Estatísticas de confiabilidade</b>” onde é apresentado o valor do Alfa de Cronbach</p>	<p>Questionários que apresentem um valor mínimo de 0,70 para este parâmetro são aceitáveis.</p> <p><b>Referência:</b> Hair Jr. et al. (2019)</p>
Avaliação da adequação da amostra	<p><b>Teste de esfericidade de Bartlett e KMO no SPSS</b></p> <p>Analisar → Redução de dimensão → Fator → Escolha as variáveis → Na aba “Descritivo” selecione “Teste de esfericidade de Bartlett e KMO” → Continuar → OK</p> <p><b>Correlações parciais da matriz anti-imagem</b></p> <p>Analisar → Redução de dimensão → Fator → Escolha as variáveis → Na aba “Descritivo” selecione “Anti-imagem” → Continuar → OK</p>	<p>Verificar a saída “<b>Teste de KMO e Bartlett</b>” onde são apresentados a medida de adequação da amostra (MSA) do teste KMO e o nível de significância do teste de esfericidade de Bartlett</p> <p>Verificar a saída “<b>Matrizes anti-imagem</b>”. O procedimento gera duas matrizes. A de baixo, “Correlação anti-imagem” tem na diagonal os MSAs parciais e nos outros pontos as correlações parciais.</p>	<p>MSA aceitável &gt; 0,50 → amostra é adequada</p> <p>Teste de esfericidade p (sig) ≤ 0,05 → amostra é adequada</p> <p><b>Referência:</b> Hair Jr. et al. (2019)</p> <p>MSAs parciais &gt; 0,5 → amostra é adequada</p> <p>Correlações parciais &lt; 0,7 → amostra é adequada</p> <p><b>Referência:</b> Hair Jr. et al. (2019)</p>
Análise da estrutura fatorial	<p><b>Determinação dos fatores com rotação</b></p> <p>Analisar → Redução de dimensão → Fator → Escolha as variáveis → Na aba “Extração” escolha o método “Fatoração pelo eixo principal” → Nesta aba, escolha também quantos fatores reter (sugestão: marcar autovalores maiores que 1 → Na aba “Rotação” escolha um método de rotação (sugestão: escolher Varimax) → Na aba “Opções” escolha “Suprimir coeficientes pequenos” (sugestão por suprimir valores abaixo de 0,30) → Continuar → OK</p>	<p>Verificar a saída “<b>Comunalidades</b>” que detalha quão bem cada variável é explicada pelo conjunto dos fatores</p> <p>Verificar a saída “<b>Variância total explicada</b>”</p> <p>Verificar a saída “<b>Matriz dos fatores rotativa</b>”</p>	<p>Se uma variável possui uma comunalidade muito baixa (&lt; 0,30) ela pode ser excluída da análise. O ideal é rodar o modelo novamente sem esta variável</p> <p>A análise desta saída ajuda a escolher quantos fatores reter. Alguns critérios para escolha do número de fatores estão listados abaixo. Na prática, um mesmo pesquisador pode utilizar mais de um critério para determinar quantos fatores devem ser extraídos.</p> <ul style="list-style-type: none"> <li>• Critério da raiz latente: escolha os fatores com autovalores maiores que 1 (Verificar a coluna “Total”).</li> <li>• Critério da porcentagem de variância explicada: em ciências humanas, uma solução que explique 60% da variância dos dados costuma ser satisfatória (Verificar a primeira coluna “% cumulativa”).</li> </ul> <p>Ao analisar esta matriz, aparecerão as respectivas cargas fatoriais de cada item em cada fator. Ao pesquisador, cabe julgar o resultado e verificar se o resultado faz sentido à luz da teoria estudada. Para julgar a significância das cargas fatoriais:</p> <ul style="list-style-type: none"> <li>• Cargas entre 0,3 e 0,4: atendem a um nível mínimo de interpretação da estrutura</li> <li>• Cargas acima de 0,5: praticamente significantes</li> <li>• Cargas acima de 0,7: meta de qualquer análise fatorial</li> </ul> <p>Após a rotação e análise das cargas, alguns problemas podem persistir como, por exemplo, uma variável não ter carga significativa; mesmo tendo carga significativa, sua comunalidade é muito baixa; ou a variável apresenta carga cruzada. Algumas soluções corretivas seriam:</p> <ul style="list-style-type: none"> <li>• ignorar a variável problemática;</li> <li>• avaliar possível eliminação da variável;</li> <li>• tentar novos modelos de rotação;</li> <li>• diminuir ou aumentar o número de fatores;</li> <li>• mudar a análise (de AFC para ACP, por exemplo).</li> </ul> <p>Por último, dá-se um nome a cada fator, lembrando que as variáveis com maior carga devem ter maior influência sobre a designação do fator.</p>

**Fonte:** Elaborado pela autora.