

Avaliação da carga de trabalho: adaptação transcultural, validade de conteúdo e confiabilidade de instrumento

Workload assessment: cross-cultural adaptation, content validity and instrument reliability

Evaluación de la carga de trabajo: adaptación transcultural, validez de contenido y confiabilidad del instrumento

Caroline Lopes Ciofi-Silva^I

ORCID: 0000-0003-0893-7255

Luciana Cordeiro^{II}

ORCID: 0000-0003-2912-1087

Naila Albertina Oliveira^{IV}

ORCID: 0000-0001-8340-5334

Giulia Marcelino Mainardi^{III}

ORCID: 0000-0001-6849-6012

Anna Sara Levin^{III}

ORCID: 0000-0003-2427-8368

Rodrigo Maximiano Antunes de Almeida^V

ORCID: 0000-0002-2663-5464

Juliana Falasco-Fantinatti^{VI}

ORCID: 0000-0002-2833-3937

Maria Clara Padoveze^{III}

ORCID: 0000-0002-1912-7293

^IUniversidade Estadual de Campinas. Campinas, São Paulo, Brasil.

^{II}Universidade Federal de Pelotas. Pelotas, Rio Grande do Sul, Brasil.

^{III}Universidade de São Paulo. São Paulo, São Paulo, Brasil.

^{IV}Universidade Federal de São Paulo. São Paulo, São Paulo, Brasil.

^VUniversidade Federal de Itajubá. Itajubá, Minas Gerais, Brasil.

^{VI}Hospital Municipal de Paulínia. Paulínia, São Paulo, Brasil.

Como citar este artigo:

Ciofi-Silva CL, Cordeiro L, Oliveira NA, Mainardi GM, Levin AS, Almeida RMA, et al. Workload assessment: cross-cultural adaptation, content validity and instrument reliability. Rev Bras Enferm. 2023;76(3):e20220556. <https://doi.org/10.1590/0034-7167-2022-0556pt>

Autor Correspondente:

Caroline Lopes Ciofi-Silva
E-mail: clciofi@unicamp.br



EDITOR CHEFE: Antonio José de Almeida Filho

EDITOR ASSOCIADO: Ana Fátima Fernandes

Submissão: 07-10-2022

Aprovação: 07-02-2023

RESUMO

Objetivos: adaptar, validar o conteúdo e avaliar a confiabilidade do instrumento *National Aeronautics and Space Administration – Task Load Index*, traduzido para o português brasileiro.

Métodos: estudo metodológico, dividido em cinco etapas: tradução; síntese; retrotradução; avaliação da versão em português por comitê de especialistas; pré-teste e validação de conteúdo da versão final por profissionais de saúde atuantes em unidades de internação. Foram calculados o Índice de Validade de Conteúdo (IVC) (mínimo 0,80) e o alfa de Cronbach (mínimo 0,70). **Resultados:** na primeira rodada, na análise de concordância da versão traduzida, três itens não alcançaram o valor mínimo do IVC. Optou-se pela remoção do enunciado. O título do instrumento e os itens “desempenho” e “esforço” foram alterados. Houve consenso e aprovação da versão final na etapa de pré-teste. **Conclusões:** o instrumento Índice NASA de carga de tarefa, adaptado para o português brasileiro, apresenta evidências de confiabilidade e validade de conteúdo.

Descritores: Carga de Trabalho; Tradução; Estudo de Validação; Pessoal de Saúde; Condições de Trabalho.

ABSTRACT

Objectives: to adapt, validate the content and assess the reliability of the instrument *National Aeronautics and Space Administration – Task Load Index*, translated into Brazilian Portuguese. **Methods:** a methodological study, divided into five steps: translation; synthesis; back-translation; assessment of the Portuguese version by an expert committee; pre-test and content validity of the final version by health professionals working in inpatient units. The Content Validity Index (CVI) (minimum 0.80) and Cronbach's alpha (minimum 0.70) were calculated. **Results:** in the first round, in the agreement analysis of the translated version, three items did not reach the minimum CVI value. It was decided to remove the statement. The instrument title and items “performance” and “effort” were changed. There was consensus and approval of the final version in the pre-test step. **Conclusions:** the NASA Task Load Index instrument, adapted to Brazilian Portuguese, presents reliability and content validity evidence. **Descriptors:** Workload; Translating; Validation Study; Health Personnel; Work.

RESUMEN

Objetivos: adaptar, validar el contenido y evaluar la confiabilidad del instrumento *National Aeronautics and Space Administration – Task Load Index*, traducido al portugués brasileiro. **Métodos:** estudio metodológico, dividido en cinco etapas: traducción; síntesis; traducción inversa; evaluación de la versión portuguesa por un comité de expertos; pretest y validación de contenido de la versión final por profesionales de la salud que actúan en unidades de hospitalización. Se calculó el Índice de Validez de Contenido (IVC) (mínimo 0,80) y el alfa de Cronbach (mínimo 0,70). **Resultados:** en la primera ronda, en el análisis de concordancia de la versión traducida, tres ítems no alcanzaron el valor mínimo de IVC. Se decidió eliminar la declaración. Se modificó el título del instrumento y los ítems “desempeño” y “esfuerzo”. Hubo consenso y aprobación de la versión final en la etapa de pre-prueba. **Conclusiones:** el instrumento *NASA Task Load Index*, adaptado al portugués brasileiro, presenta evidencias de confiabilidad y validez de contenido.

Descriptores: Carga de Trabajo; Traducción; Estudio de Validación; Personal de Salud; Condiciones de Trabajo.

INTRODUÇÃO

O universo do trabalho é objeto de estudo há centenas de anos, sendo que o conhecimento acerca da carga de trabalho, tanto em relação à definição do constructo quanto ao seu impacto na psicodinâmica dos trabalhadores, vem sendo acumulado nas últimas décadas. O processo de globalização e as mudanças no mundo do trabalho, considerando a revolução tecnológica e a entrada do capital estrangeiro na economia, trazem ainda maior complexidade para esse campo⁽¹⁾.

As novas formas e relações de trabalho têm agravado os problemas dele decorrentes e impactado fortemente física e psicologicamente os trabalhadores⁽¹⁾. Assim, a discussão sobre carga de trabalho, termo complexo e polissêmico, e suas possibilidades de mensuração, permanece atual em todo o planeta.

Carga de trabalho é um termo amplo e multifacetado que engloba o “custo” de realizar uma tarefa por um operador humano (por exemplo, fadiga, estresse, adoecimento, acidentes, etc.)⁽²⁾. Hart e Staveland (1988), considerando tanto a forma como as pessoas formularam opiniões acerca de carga de trabalho quanto a expressão subjetiva da questão, isolaram e definiram fatores relevantes para avaliação desse constructo⁽²⁻³⁾.

A maior parte das escalas que se propõem a medir carga de trabalho ou carga de tarefa consideram a subjetividade dos respondentes, sendo necessário corrigir a alta variabilidade das respostas dos trabalhadores; outro problema seria a variação da carga de trabalho, a depender da tarefa desempenhada e área de atuação profissional. Por conseguinte, um índice que considera a multidimensionalidade do fenômeno e, ao mesmo tempo, reduz a variabilidade subjetiva, a fim de computar a carga de trabalho global, foi proposto⁽³⁾.

Atualmente, o *National Aeronautics and Space Administration-Task Load Index* (NASA-TLX) é um índice de domínio público. Originalmente desenvolvido em inglês e inicialmente voltado para o campo da aviação, o índice avalia a carga de trabalho durante ou após a realização de uma tarefa. O NASA-TLX pode ser aplicado ao longo ou imediatamente após a realização de uma tarefa ou conjunto de tarefas, propondo-se a avaliar o desempenho em seis dimensões, a fim de determinar a carga de trabalho global de uma tarefa. Três das seis dimensões se referem a aspectos ou exigências impostas pelo sujeito (mental, física e temporal), e as outras três, à interação sujeito-tarefa (esforço, frustração e realização)⁽³⁾.

Quanto às seis dimensões do NASA-TLX, a dimensão “demanda mental” diz respeito ao raciocínio, tomada de decisão ou cálculo necessários para realizar a tarefa; a dimensão “demanda física” mensura a quantidade e intensidade de atividade física que é requerida para realizar a tarefa; a “demanda temporal” está relacionada à quantidade de tempo envolvida para completar a tarefa; a dimensão “esforço” avalia o quanto o trabalhador precisa se esforçar para realizar a tarefa; a dimensão “desempenho” diz respeito à realização da tarefa de forma satisfatória; e, finalmente, a dimensão “frustração” avalia sentimentos do participante durante a realização da tarefa⁽³⁾. A escala de resposta é apresentada por meio de uma linha horizontal, dividida em linhas verticais em 20 intervalos iguais, sem demarcações numéricas, com extremidades bipolares, variando entre zero (baixa carga de trabalho) e 100 (alta carga de trabalho). Recomenda-se assinalar a localização desejada para cada dimensão, baseada na experiência da tarefa executada.

O resultado é obtido analisando o ponto correspondente e acrescentando incrementos de 5 a cada demarcação. Quanto à interpretação dos resultados, os escores podem ser obtidos de cada dimensão ou pode ser calculada a média, obtendo-se a carga de trabalho global. Até o momento, não há evidências fortes de valores de referência e classificações dos escores. No entanto, sugere-se que sejam realizadas comparações de resultados do instrumento relativos a tarefas semelhantes⁽³⁻⁴⁾.

O NASA-TLX é um dos índices mais utilizados em todo o mundo para medir carga de trabalho⁽⁴⁾. Vinte anos após seu desenvolvimento, foi traduzido para mais de doze idiomas, aplicado em diferentes campos e modificado de diversas formas⁽²⁾. Uma versão *online* para administração do índice⁽⁵⁾ e um *software* para simplificar os processos de coleta, processamento e armazenamento de dados⁽⁶⁾ foram também criados. O índice foi submetido a avaliações em termos de confiabilidade e sensibilidade⁽⁷⁾, sendo comparado com outros instrumentos de mensuração de carga de trabalho⁽⁷⁻⁸⁾.

Pesquisas brasileiras têm se dedicado a medir carga de trabalho utilizando, principalmente instrumentos que detectam sofrimento mental, que avaliam qualidade de vida ou que dimensionam fatores de risco psicossociais no trabalho⁽⁹⁾. O NASA-TLX tem sido utilizado no Brasil em investigações de diversos contextos, com destaque para a indústria, para estudantes de graduação, pós-graduação e profissionais da área da saúde⁽¹⁰⁾. No contexto internacional, esse índice também foi utilizado em estudo com simulação clínica para avaliação de equipamentos de proteção individual (EPI)⁽¹¹⁾.

O contexto da pandemia do coronavírus trouxe questões importantes acerca dos processos de trabalho e das condições de saúde dos trabalhadores da saúde, considerados trabalhadores essenciais. Soma-se aos problemas crônicos da saúde pública brasileira, tais como o sub-financiamento do Sistema Único de Saúde (SUS), o congelamento dos gastos no setor, a deterioração dos serviços, a precarização das condições de trabalho e os desafios emergentes. Problemas ligados à gestão do trabalho e capacitação de pessoal diante de abertura de novos leitos na alta complexidade requereram reorganização do processo de trabalho, trazendo novas exigências aos trabalhadores. Nesse sentido, a mensuração da carga de trabalho pode ser especialmente útil neste contexto, a fim de propor ações para minimizar problemas advindos do trabalho⁽¹²⁾.

Ainda que já utilizado no Brasil^(10,13), não há registro de adaptação transcultural e validação para o português brasileiro do NASA-TLX. Partindo dessa realidade, aponta-se o risco de traduções e modificações de escalas desenvolvidas em idioma estrangeiro sem estudo aprofundado de sua validade e precisão de avaliações efetuadas, podendo comprometer a equivalência entre a escala original e a versão traduzida e, conseqüentemente, suas características e domínios avaliados. Para tanto, a utilização de um método rigoroso de adaptação transcultural se faz necessário antes de se aplicar um instrumento desenvolvido em outro idioma e cultura⁽¹⁴⁾.

OBJETIVOS

Realizar a adaptação transcultural, validação de conteúdo e avaliação da confiabilidade do NASA-TLX para o português

brasileiro, com vistas a garantir que o índice seja adequado para aplicação em estudos no Brasil em diferentes áreas de atuação profissional.

MÉTODOS

Aspectos éticos

O estudo foi conduzido de acordo com as diretrizes de ética nacionais e internacionais, sendo aprovado pelo Comitê de Ética em Pesquisa da Escola de Enfermagem da Universidade de São Paulo, cujo parecer está anexado à presente submissão. O Termo de Consentimento Livre e Esclarecido foi assinado por todos os indivíduos envolvidos no estudo: de forma *online*, para os especialistas, e de forma escrita, para os profissionais de saúde.

Desenho do estudo

Trata-se de um estudo metodológico de tradução e adaptação transcultural. Foi realizado contato por meio de correio eletrônico com os autores da versão original do instrumento, solicitando autorização para realização da adaptação transcultural. Os autores informaram que, por ser de domínio público, modificações e traduções estão autorizadas.

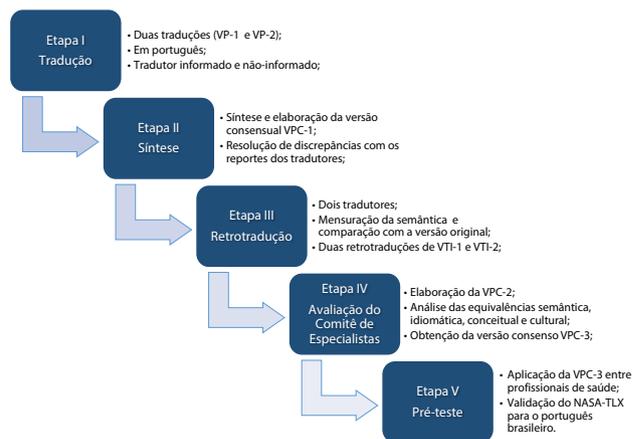
Protocolo do estudo

O processo de adaptação transcultural e validação de conteúdo foi baseado no método proposto por Beaton e em outros estudos de adaptação transcultural de instrumentos psicométricos realizados no cenário brasileiro⁽¹⁴⁻¹⁷⁾. As etapas metodológicas foram divididas em cinco etapas, ilustradas na Figura 1 e detalhadas a seguir:

- I. Tradução para a língua portuguesa: realizada por dois tradutores bilíngues, atuantes em empresas certificadas em tradução, de forma independente, sendo que apenas um dos tradutores possuía conhecimento do objetivo do instrumento (tradutor informado), resultando em duas versões em português (VP-1 e VP-2);
- II. Síntese da tradução e elaboração de uma versão consensual na língua portuguesa: realizada por um terceiro tradutor que, de forma neutra, realizou um relatório demonstrando os meios utilizados para resolver as divergências. As diferenças foram analisadas e discutidas entre os três tradutores, resultando na versão consensual em português (VPC-1);
- III. Retrotradução: dois tradutores, com certificação na língua inglesa, realizaram a tradução da versão consensual da língua portuguesa para a língua inglesa, resultando em duas novas versões (VTI-1 e VTI-2). Esses tradutores não tiveram acesso à versão original e conceitos abordados;
- IV. Avaliação pelo comitê de especialistas: obtidas as versões após a retrotradução, os pesquisadores principais do projeto elaboraram uma nova versão consensual em português (VPC-2). Essa versão, juntamente com as demais (original, VP-1, VP-2, VPC-1, VT-1, VT-2), foram dispostas em um quadro e enviadas para avaliação por um comitê de

especialistas. Esse comitê foi composto por profissionais especialistas com, no mínimo, titulação de mestre, com conhecimento fluente na língua inglesa, de diferentes áreas de conhecimento e atuação, a fim de obter uma avaliação multidisciplinar do processo. Além disso, foram convidados profissionais atuantes nas cinco regiões do Brasil, visando dirimir possíveis disparidades culturais no processo de tradução. Os especialistas foram selecionados por meio de análise do *Curriculum Lattes*, sendo a amostra por conveniência, e foram convidados para participar por meio de correio eletrônico. Após o aceite do convite, os especialistas foram orientados a registrarem seus pareceres, respondendo em uma escala do tipo Likert⁽¹⁸⁾ de quatro pontos (“concordo totalmente”, “concordo”, “discordo” e “discordo totalmente”), se concordavam ou não em relação às equivalências semântica, idiomática, conceitual e cultural de cada item da versão traduzida comparada à original. Adicionalmente, para cada item, havia um campo para escreverem sugestões de alterações ou comentários. A partir da técnica Delphi⁽¹⁹⁾, foi obtida a versão consenso, sendo nomeada VPC-3⁽¹⁸⁻¹⁹⁾;

- V. Pré-teste: a VPC-3 foi aplicada em profissionais da saúde de um hospital público de uma cidade do interior do estado de São Paulo. Os critérios de inclusão dos participantes dessa etapa foram ser profissionais com experiência na assistência a pacientes em internação hospitalar há pelo menos seis meses. Foram excluídos os profissionais com alguma restrição física ou psicológica para o trabalho, documentada junto ao serviço de medicina de trabalho da instituição. Os profissionais selecionados para amostra foram orientados a preencher os seguintes formulários impressos no final dos respectivos turnos de trabalho: VPC-3 referente ao conjunto de tarefas realizadas no turno de trabalho; um questionário com dados sociodemográficos; e um questionário com perguntas sobre a avaliação geral e compreensão do conteúdo da VPC-3. A coleta de dados da etapa de pré-teste ocorreu entre os meses de janeiro e fevereiro de 2022.



Fonte: adaptado de Beaton et al., 2000.

Figura 1 - Fluxograma do processo de adaptação transcultural e validação do NASA-TLX

Análise dos resultados e estatística

Os dados obtidos de todas as etapas foram registrados e organizados em planilhas do *Microsoft Excel for Windows*, possibilitando a análise descritiva dos resultados. Na etapa da avaliação do comitê de especialistas, foi realizado o teste de Índice de Validade de Conteúdo (IVC) para avaliar a concordância entre os mesmos. O valor mínimo aceitável para o IVC foi de 0,80⁽²⁰⁾. Na etapa de pré-teste, foram realizados os testes Alfa (α) de Cronbach para atestar a confiabilidade do questionário⁽²¹⁾, além do IVC no questionário sobre a avaliação geral e da compreensão da versão traduzida do NASA-TLX. Os testes estatísticos foram realizados por meio do programa *Statistical Package for the Social Sciences* (SPSS), versão 26.0. Na análise estatística, o nível de significância adotado foi de 95%.

RESULTADOS

As três primeiras etapas do processo de adaptação transcultural e validação do NASA-TLX transcorreram com auxílio de empresas certificadas em tradução de documentos na língua inglesa. Pequenas discrepâncias entre palavras sinônimas e graus dos adjetivos empregados na composição das sentenças foram resolvidas pelos pesquisadores principais.

Na etapa IV, trinta especialistas foram convidados para compor o comitê, no entanto 14 aceitaram o convite e responderam os formulários. Onze especialistas eram da região Sudeste do Brasil (sete do estado de São Paulo e um de Minas Gerais); um, da região Norte (estado do Amazonas); um, da região Centro-Oeste (Distrito Federal); e um era brasileiro que residia nos Estados Unidos da América no período do estudo. Quatro eram profissionais atuantes na área de exatas (três da engenharia e um da física aplicada); três, na área de humanas (letras), com ênfase na língua inglesa; e sete, na área de biológicas (cinco da enfermagem, um da nutrição e um da terapia ocupacional). Um participante era livre-docente; oito possuíam titulação máxima de doutor; e cinco possuíam titulação máxima de mestre.

Na análise de concordância da versão traduzida do NASA-TLX, os itens “título e enunciado”, “desempenho” e “esforço”, nos quesitos das avaliações semântica, idiomática e conceitual, não alcançaram o valor mínimo estabelecido do IVC=0,80 (Tabela 1); portanto, foram reformulados e encaminhados para a segunda etapa de avaliação pelo comitê de especialistas. Dos 14 especialistas participantes na primeira rodada de consenso, apenas sete participaram da segunda rodada de consenso, mesmo após três tentativas de contato por meio de correio eletrônico.

Tabela 1 - Análise dos Índices De Validação De Conteúdo da versão consenso em português - 2 (VPC-2), a partir das respostas dos especialistas na primeira rodada de consenso, São Paulo, São Paulo, Brasil, 2022

	Semântica	Idiomática	Cultural	Conceitual
Título e enunciado	0,714	0,714	0,786	0,714
Demanda mental	0,929	1,000	0,929	0,929
Demanda física	0,929	1,000	0,929	0,857
Demanda temporal	0,929	1,000	1,000	0,857
Desempenho	0,714	0,857	0,929	0,571
Esforço	0,786	0,643	0,929	0,500
Frustração	0,929	0,929	1,000	0,929

Sobre o item “título e enunciado”, quanto ao título do instrumento na VPC-2, nove especialistas (64,28%) sugeriram adotar o termo “carga de trabalho” ao invés de “carga de tarefa”. Após a análise desse resultado, os pesquisadores principais do estudo se reuniram com dois dos tradutores envolvidos no processo de tradução e decidiram manter a tradução do termo original “task” para “tarefa”. Quanto ao enunciado, após terem sido identificadas inconsistências entre diferentes versões do instrumento, foi realizada uma consulta com os autores do instrumento original, os quais relataram que o enunciado não estava contemplado na elaboração inicial do NASA-TLX. Dessa forma, optou-se pela remoção do enunciado na versão do NASA-TLX traduzido para o português brasileiro, uma vez que não se trata de instruções para preenchimento da escala, sendo, portanto, prescindível. Essas decisões foram apresentadas aos especialistas na segunda rodada da técnica Delphi⁽¹⁹⁾, juntamente com as devidas justificativas, obtendo-se consenso para ambas. Dessa forma, a versão em português brasileiro manteve o título do instrumento como “Índice NASA de carga de tarefa” na VPC-3.

Ainda na primeira etapa de análise do comitê de especialistas, não foi obtido o consenso a respeito dos termos utilizados nos extremos da escala de resposta para o item “desempenho”, sendo, portanto, alterados de “perfeito” e “fracasso”, conforme proposto na VPC-2, para “muitíssimo” e “pouquíssimo”, mantendo o mesmo formato para as outras cinco perguntas da escala na VPC-3. Sobre o item “esforço”, o consenso não foi obtido referente à pergunta, a qual foi modificada de “o quão difícil foi realizar a tarefa para atingir o seu nível de desempenho?” (de acordo com a versão VPC-2) para “quanto esforço foi necessário para atingir o seu nível de desempenho no trabalho?”.

Obtido o consenso entre os especialistas também para esses itens, a VPC-3 foi aplicada conforme descrito na etapa do pré-teste. Cinquenta profissionais de saúde de diferentes categorias foram incluídos, e as características sociodemográficas dos participantes estão apresentadas na Tabela 2. Quanto às respostas dos participantes para a versão traduzida para o português da VPC-3, as médias obtidas foram: “demanda mental” = 74,21; “demanda física” = 68,82; “demanda temporal” = 71,96; “desempenho” = 24,31; “esforço” = 73,82; “frustração” = 52,25.

No que diz respeito à avaliação geral da VPC-3, 38 (76%) participantes consideraram-na boa, nove (18%), regular, e três (6%), ruim. Quarenta e três participantes (86%) relataram que todas as questões foram fáceis de entender; 13 (24%) responderam que algumas questões foram difíceis de entender; e 2 (4%) participantes responderam que todas as questões foram difíceis de entender. Sobre a compreensão da escala de respostas do instrumento, 36 (72%) participantes responderam que foi fácil; 12 (24%) consideraram a compreensão regular; e 2 (4%) consideraram difícil. Para a validação da compreensão de conteúdo das seis perguntas individuais da VPC-3, os valores de IVC obtidos foram superiores a 0,80 na etapa do pré-teste. A consistência interna do instrumento foi confirmada por meio do resultado do alfa de Cronbach total = 0,757. Na Tabela 3, estão apresentados os valores de IVC para cada item, bem como os escores de alfa de Cronbach, caso o item fosse excluído.

Em relação às sugestões e comentários descritos pelos participantes do pré-teste, de forma geral, os participantes destacaram que o instrumento mensura a carga de trabalho de um único turno de trabalho e que oscilações de demandas de trabalho

podem ocorrer entre os turnos, influenciando no resultado do instrumento. Adicionalmente, comentou-se que o questionário não aborda a influência do trabalho em equipe na carga de trabalho, e a inclusão de diferentes sentimentos em uma única questão na dimensão “frustração” pode dificultar a escolha da graduação na escala de resposta.

Tabela 2 - Características sociodemográficas dos profissionais de saúde que participaram da etapa de pré-teste, São Paulo, São Paulo, Brasil, 2022

Características	n	%
Sexo		
Feminino	39	78
Masculino	11	22
Profissão		
Enfermeiro	12	24
Técnico em enfermagem	26	52
Auxiliar de enfermagem	3	6
Médico	5	10
Fisioterapeuta	4	8
Local de trabalho		
Enf. cirúrgica	8	16
Enf. clínica médica	10	20
Enf. de ginecologia obstetrícia	6	12
Enf. pediátrica	8	16
Unidade de Terapia Intensiva adulto	15	30
Vários*	3	6
Experiência profissional		
até 1 ano	4	8
1-5 anos	29	58
6-10 anos	10	20
11-20 anos	6	12
mais de 20 anos	1	2
Duração do turno de trabalho		
12 horas	44	88
11 horas	1	2
6 horas	3	6
24 horas	2	4
Turno de trabalho		
Diurno	28	56
Noturno	22	44

Enf. – enfermagem; Vários* – profissionais que atendiam diferentes setores no mesmo turno de trabalho.

Tabela 3 - Análise da validação de conteúdo e consistência interna da VPC-3 na etapa de pré-teste, São Paulo, São Paulo, Brasil, 2022

	IVC	α de Cronbach, se o item for excluído
Demanda mental	0,900	0,727
Demanda físico	0,900	0,694
Demanda temporal	0,840	0,678
Desempenho	0,860	0,722
Esforço	0,920	0,706
Frustração	0,860	0,727

IVC – Índice de Validação do Conteúdo.

DISCUSSÃO

O presente artigo teve como objetivo descrever o processo de tradução e adaptação transcultural do instrumento NASA-TLX para avaliação da carga de trabalho. Percorridas todas as etapas metodológicas, com rigor e de acordo com as etapas fundamentadas na literatura, foram alcançados os resultados pretendidos. Obteve-se como resultado que a versão consenso em português brasileiro – 3 (VPC-3) é válida e confiável, sendo nomeada “Índice

NASA de carga de tarefa” (o arquivo pode ser acessado por meio do material suplementar).

Por constituir-se um método rigoroso, que requer a destinação de recursos e esforços dos pesquisadores, a escolha de conduzir o processo de tradução e adaptação transcultural de um instrumento deve priorizar instrumentos relevantes e que atendam ao objetivo para o qual foram elaborados. A justificativa da escolha pelo NASA-TLX foi fundamentada por ser um instrumento curto, de fácil aplicação e que pode ser utilizado em diferentes contextos de trabalho, não restritos ao campo de trabalho da saúde, além de já ser amplamente utilizado em diferentes países⁽²²⁾.

Nos estágios iniciais do processo de tradução, realizados por tradutores profissionais, foram evidenciadas divergências entre a tradução de termos específicos, principalmente quando o tradutor desconhecia o objetivo do instrumento. A depender do termo traduzido, diferentes interpretações podem ser realizadas, variando de acordo com o contexto cultural da região. Nesse sentido, destaca-se a importância da inclusão do critério “equivalência cultural” na etapa de avaliação do comitê de especialistas neste estudo. Alguns dos especialistas apontaram que a tradução literal dos termos “perfect” (perfeito) and “failure” (fracasso) da escala de resposta original do item “performance” do NASA-TLX poderia ser interpretada equivocadamente pelos usuários do instrumento. Portanto, esses termos foram substituídos, mantendo os termos das escalas de respostas dos demais itens: “pouquíssimo” e “muitíssimo”. Na adaptação transcultural, esse tipo de substituição facilita a interpretação da escala, fazendo com que os termos utilizados sejam mais familiares e acessíveis à população-alvo⁽²²⁾. Os termos “good” e “poor” também são encontrados em outras versões do instrumento⁽²⁾.

Dadas as diferenças regionais e culturais do Brasil, no intuito de assegurar a confiabilidade da versão traduzida e adaptada em todo o território brasileiro, foram convidados profissionais com diferentes áreas de atuação e das cinco regiões do país para comporem o comitê de especialistas. No entanto, profissionais de apenas três regiões do Brasil responderam o convite e enviaram suas respostas, com predomínio de profissionais que residiam na região Sudeste. Não obstante a ampliação do uso das tecnologias digitais com a pandemia de COVID-19, facilitando os contatos a despeito da distância física, identificou-se o baixo retorno das respostas. A dificuldade em obter respostas de especialistas é considerada uma das desvantagens da técnica Delphi de consenso⁽²³⁾.

Quanto ao título do instrumento “Índice NASA de carga de tarefa”, não foi obtido consenso também com relação à utilização da expressão “carga de tarefa” ao invés de “carga de trabalho”. Na versão final do instrumento, obteve-se o consenso em manter a tradução do termo original “task” como “tarefa”, considerando que o questionário poderá ser aplicado tanto em relação a uma tarefa específica quanto a um conjunto de tarefas após um turno de trabalho, como realizado na etapa de pré-teste deste estudo. A utilização do NASA-TLX foi validada para avaliação da carga de um conjunto de tarefas por meio de um estudo envolvendo adultos com diabetes do tipo 1, nos Estados Unidos, após a realização de um conjunto de atividades de autocuidado⁽²⁴⁾.

Em outro estudo, o NASA-TLX também foi utilizado para avaliar a carga de um conjunto de tarefas entre enfermeiros atuantes em Unidades de Cuidados Críticos após seus turnos de

trabalho⁽²⁵⁾. Portanto, há subsídios para aplicação do NASA-TLX traduzido para o português brasileiro após a realização de uma tarefa específica ou um conjunto de tarefas, sendo possível adotar também o termo “carga de trabalho”. Entretanto, sugere-se que, na aplicação do NASA-TLX após um conjunto de tarefas ou um turno de trabalho, os pesquisadores considerem que diferentes fatores podem influenciar nos resultados que serão obtidos. É provável que a carga de trabalho não seja um elemento isolado, mas sim compreendido no conjunto com outras atividades em um determinado período. Assim sendo, é desejável que sejam realizados estudos adicionais analisando se há diferenças na acurácia do NASA-TLX no uso específico de uma determinada tarefa ou no conjunto de tarefas ao longo de um turno de trabalho.

Na etapa de pré-teste, mais de 70% dos profissionais de saúde participantes avaliaram que, de forma geral, a versão validada para o português brasileiro do NASA-TLX era boa e que tanto as questões compreendidas no instrumento quanto a escala de resposta eram fáceis de entender. Adicionalmente, a confiabilidade e a consistência foram atestadas por meio dos escores acima de 0,80 do IVC e alfa de Cronbach acima de 0,70. Esses dados corroboram que o NASA-TLX adaptado para o português brasileiro é útil para avaliação de carga de tarefa no contexto da assistência à saúde, principalmente diante de crises sanitárias, como a pandemia de COVID-19, o que pode influenciar na qualidade da assistência⁽²⁶⁾.

Ao longo das décadas desde a sua elaboração, a análise dos escores do NASA-TLX sofreu adaptações, e uma delas consiste na utilização dos escores brutos de cada domínio, além de um escore geral para estimar a carga de tarefa global, por meio da soma dos escores individuais. Anteriormente, os autores recomendavam a atribuição de pesos a cada domínio, conforme o número de vezes que cada um deles era avaliado como significativo pelos participantes. No entanto, esse processo tem sido considerado complexo e ineficaz. A utilização dos escores brutos, como realizada no pré-teste deste estudo, e frequentemente adotada em outros estudos, facilita a aplicação do instrumento e possibilita comparar os escores individuais de cada domínio, propiciando a inferência de qual domínio exerce maior ou menor influência na carga de tarefa⁽²⁶⁻²⁷⁾.

Shoja *et al.* evidenciaram, por meio da aplicação do NASA-TLX entre diferentes profissionais da saúde responsáveis pelo atendimento de pacientes com COVID-19, que os enfermeiros relataram maiores escores de carga de tarefa, e a dimensão “demanda mental” foi a que apresentou maiores escores⁽²⁸⁾. Na etapa de pré-teste, a dimensão “demanda mental” também foi a que apresentou maior média de escore. No entanto, por se tratar de um estudo de adaptação transcultural e validação em uma amostra restrita, não é possível generalizar os resultados obtidos sobre a aplicação da escala.

Limitações do estudo

Este estudo possui algumas limitações. O baixo índice de respostas, com participação de três dentre as cinco regiões brasileiras, pode ter limitado a contextualização cultural da tradução. Além disso, a tradução realizada foi feita para o português brasileiro e pode não ter plena correspondência para o português de outros países. A realização da etapa de pré-teste em um único serviço de saúde permitiu a análise dentro de um cenário específico,

podendo não necessariamente corresponder a outros cenários similares. Não obstante, essas limitações não comprometem o resultado do processo de adaptação transcultural, validação de conteúdo e confiabilidade do NASA-TLX, que se apresenta como um instrumento disponível para uso imediato no contexto brasileiro.

Contribuições para as áreas da enfermagem e saúde

No campo da saúde, tendo em vista perspectivas de futuras pesquisas, a versão do NASA-TLX em português brasileiro poderá ser utilizada também para comparar diferentes dispositivos de trabalho, tais como novos EPI, e ainda na implementação de novas tecnologias de processos de trabalho, tais como aplicativos computadorizados. Adicionalmente, esse instrumento poderá servir de parâmetro de comparação de carga de tarefa, auxiliando no processo decisório de planejamento e alocação de equipes de trabalho. Novos estudos poderão também ampliar as propriedades psicométricas do instrumento aplicáveis a outros tipos de atividades de trabalho.

CONCLUSÕES

O instrumento “Índice NASA de carga de tarefa”, traduzido e adaptado para o português brasileiro, apresenta evidências de confiabilidade e validade de conteúdo. Apesar de o instrumento original possuir poucos itens, identificou-se que as traduções de alguns termos poderiam ser passíveis de interpretação equivocada, comprometendo a confiabilidade do instrumento. Após as etapas de adaptação, essas falhas foram minimizadas, resultando em uma ferramenta que poderá ser utilizada em diversos cenários de trabalho. Sugere-se que novos estudos psicométricos, utilizando o instrumento apresentado neste artigo, sejam realizados em outros contextos de trabalho no território brasileiro.

FOMENTO

O presente estudo foi um desdobramento do estudo multicêntrico internacional intitulado: “*Perceived workload using separate (FFP2 Respirator and Face Shield) and Integrated (Powered Air Purifying Respirator) personal protective equipment for respiratory protection*”, o qual obteve apoio financeiro da Organização Mundial da Saúde.

AGRADECIMENTO

Os autores agradecem aos pesquisadores Ying Ling Lin, da Organização Mundial da Saúde, Amy Price e Larry Fu-Nien Chu, da Universidade de Stanford (Estados Unidos da América), por fornecerem orientações sobre métodos de utilização do instrumento NASA-TLX e de análise dos resultados. E ainda, os autores agradecem à enfermeira Geowanna Aparecida Higino Silva Santos, do Hospital Municipal de Paulínia, por colaborar na viabilização da coleta de dados da etapa de pré-teste.

CONTRIBUIÇÕES

Ciofi-Silva CL, Cordeiro L, Oliveira NA, Mainardi GM e Padoveze MC contribuíram com a concepção ou desenho do estudo/

pesquisa. Ciofi-Silva CL, Cordeiro L, Oliveira NA, Levin AS, Almeida RMA, Falasco-Fantinatti e Padoveze MC contribuíram com a análise e/ou interpretação dos dados. Ciofi-Silva CL, Cordeiro L, Oliveira NA, Levin AS, Mainardi GM, Almeida RMA e Padoveze MC contribuíram com a revisão final com participação crítica e intelectual no manuscrito.

REFERÊNCIAS

1. Antunes R. Coronavírus: O trabalho sob fogo cruzado[Ebook]. São Paulo: Editora Boitempo; 2020. 52p.
2. Hart SG. NASA-Task Load Index (NASA-TLX); 20 years later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting.2006;50:904–908. <https://doi.org/10.1177/154193120605000909>
3. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Adv Psychol*. 1988;52(C):139–83. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
4. Grier RA. How high is high? a meta-analysis of NASA-TLX Global Workload Scores. Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting.2015;59(1):1727-31. <https://doi.org/10.1177/1541931215591373>
5. Sharek, D. A useable, online NASA-TLX tool. Proceedings of the Human Factors and Ergonomics Society Annual Meeting.2011;55:1375-79. <https://doi.org/10.1177/1071181311551286>
6. Cao A, Chintamani KK, Pandya AK, Ellis RD. NASA TLX: Software for assessing subjective mental workload. *Behav Res Methods*. 2009;41(1):113–7. <https://doi.org/10.3758/BRM.41.1.113>
7. Hill SG, Iavecchia HP, Byers JC, Bittner AC, Zaklad AL, Christ RE. Comparison of four subjective workload rating scales. *Hum Factors*. 1992;34(4):429–39. <https://doi.org/10.1177/001872089203400405>
8. Rubio S, Díaz E, Martín J, Puente JM. Evaluation of subjective mental workload: a comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychol*. 2004;53(1):61–86. <https://doi.org/10.1111/j.1464-0597.2004.00161.x>
9. Gruginski BE, Hembecker PK, Poletto A, Gontijo LA. Aplicação De Instrumentos De Medida De Carga Mental De Trabalho Em Publicações Recentes No Brasil. In: XXXVII Encontro Nacional De Engenharia De Producao [Internet]. 2017[cited 2002 Oct 05]. Joinville, Enegep. Available from: <https://pdfs.semanticscholar.org/d861/6712dbde39cc4aba2397537e40c0fcd40081.pdf>
10. Benin FMC, Pessa SLR. Método NASA TLX: Revisão sistemática da produção científica nacional no período de 2005 a 2017. In: IX Congresso Brasileiro De Engenharia De Produção [Internet]. 2019[cited 2002 Oct 05]. Ponta Grossa, Associação Paranaense de Engenharia de Produção. Available from: http://aprepro.org.br/conbrepro/2019/anais/arquivos/09302019_140906_5d92385644e9e.pdf
11. Drews FA, Mulvey D, Stratford K, Samore MH, Mayer J. Evaluation of a Redesigned Personal Protective Equipment Gown. *Clin Infect Dis*. 2019;69(Suppl 3):S199–205. <https://doi.org/10.1093/cid/ciz520>
12. Teixeira CFS, Soares CM, Souza EA, Lisboa ES, Pinto ICM, Andrade LR, et al. A saúde dos profissionais de saúde no enfrentamento da pandemia de Covid-19. *Cien Saude Colet*. 2020;25(9):3465–74. <https://doi.org/10.1590/1413-81232020259.19562020>
13. Dal Sasso GM, Barra DC. Cognitive Workload of Computerized Nursing Process in Intensive Care Units. *Comput Inform Nurs*. 2015;33(8):339-E1. <https://doi.org/10.1097/CIN.000000000000157>
14. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186-3191. <https://doi.org/10.1097/00007632-200012150-00014>
15. Minosso KC, Toso BRGO. Transcultural validation of an instrument to evaluate Advanced Nursing Practice competences in Brazil. *Rev Bras Enferm*. 2021;74:e20210165. <https://doi.org/10.1590/0034-7167-2021-0165>
16. Cunha CM, Nepomuceno E, Manzato RO, Cunha DCPT, Silva D, Dantas RAS. Cultural adaptation and validation of the Brazilian Version of the Patient Activation Measure-22 items. *Rev Bras Enferm*. 2018;71(4):1891–8. <https://doi.org/10.1590/0034-7167-2016-0470>
17. Freitas NO, Caltran MP, Dantas RAS, Rossi LA. Translation and cultural adaptation of the Perceived Stigmatization Questionnaire for burn victims in Brazil. *Rev Esc Enferm USP*. 2014;48(1):25-33. <https://doi.org/10.1590/S0080-623420140000100003>
18. Likert R. A technique for the measurement of attitudes. *Arch Psychol* [Internet].1932[cited 2002 Oct 05];140:5-55. Available from: https://legacy.voteview.com/pdf/Likert_1932.pdf
19. Beiderbeck D, Frevel N, von der Gracht HA, Schmidt SL, Schweitzer VM. Preparing, conducting, and analyzing Delphi surveys: Cross-disciplinary practices, new directions, and advancements. *MethodsX*. 2021;8:101401. <https://doi.org/10.1016/j.mex.2021.101401>
20. Davis LL. Instrument review: getting the most from a panel of experts. 1992;5(4):194–97. [https://doi.org/10.1016/s0897-1897\(05\)80008-4](https://doi.org/10.1016/s0897-1897(05)80008-4)
21. Cronbach LJ. Coefficient alpha and the internal structure of test. *Psychometrika*.1951;16(3):297-334. <https://doi.org/10.1007/BF02310555>
22. Coster WJ, Mancini MC. Recomendações para a tradução e adaptação transcultural de instrumentos para a pesquisa e a prática em Terapia Ocupacional. *Rev Ter Ocup Univ São Paulo*. 2015;26(1):50-7. <https://doi.org/10.11606/issn.2238-6149.v26i1p50-57>
23. Scarparo AF, Laus AM, Azevedo ALCS, Freitas MRI, Gabriel CS, Chaves LDP. Reflexões sobre o uso da Técnica Delphi em pesquisas na Enfermagem. *Rev Rene* [Internet]. 2012[cited 2002 Oct 05];13(1):242-51. Available from: <http://www.periodicos.ufc.br/rene/article/view/3803>

24. Hernandez R, Roll SC, Jin H, Schneider S, Pyatak EA. Validation of the National Aeronautics and Space Administration Task Load Index (NASA-TLX) adapted for the whole day repeated measures context. *Ergonomics*. 2022;65(7):960-975. <https://doi.org/10.1080/00140139.2021.2006317>
 25. Tubbs-Cooley HL, Mara CA, Carle AC, Gurses AP. The NASA Task Load Index as a measure of overall workload among neonatal, paediatric and adult intensive care nurses. *Intensive Crit Care Nurs*. 2018;46:64-69. <https://doi.org/10.1016/j.iccn.2018.01.004>
 26. Hertzum M. How Demanding Is Healthcare Work? a Meta-Analytic Review of TLX Scores. *Stud Health Technol Inform*. 2021;286:55-59. <https://doi.org/10.3233/SHTI210637>
 27. Hertzum M. Reference values and subscale patterns for the task load index (TLX): a meta-analytic review. 2021;64(7):869-78. <https://doi.org/10.1080/00140139.2021.1876927>
 28. Shoja E, Aghamohammadi V, Bazayar H, Moghaddam HR, Nasiri K, Dashti M, et al. Covid-19 effects on the workload of Iranian healthcare workers. *BMC Public Health*. 2020;20(1):1636. <https://doi.org/10.1186/s12889-020-09743-w>
-