


# Validação psicométrica do instrumento *Creighton* para avaliação de competências clínicas em simulação

Psychometric validation of the Creighton Competency Evaluation Instrument in simulation  
Validación psicométrica del instrumento Creighton para evaluar competencias clínicas en simulacro

Jéssica de Oliveira Veloso Vilarinho<sup>1</sup>  <https://orcid.org/0000-0003-0503-7104>

Jorge Vinícius Cestari Felix<sup>1</sup>  <https://orcid.org/0000-0002-0086-674X>

Luciana Puchalski Kalinke<sup>1</sup>  <https://orcid.org/0000-0003-4868-8193>

Alessandra Mazzo<sup>2</sup>  <https://orcid.org/0000-0001-5074-8939>

Francisco Diego Negrão Lopes Neto<sup>1</sup>  <https://orcid.org/0000-0003-2094-5928>

Radamés Boostel<sup>1</sup>  <https://orcid.org/0000-0003-2678-0001>

Nilton Orlando da Silva<sup>1</sup>  <https://orcid.org/0000-0003-1956-7376>

Amanda Carolina de Oliveira Bialecki Fontoura<sup>1</sup>  <https://orcid.org/0000-0002-4983-7915>

## Como citar:

Vilarinho JO, Feliz JV, Kalinke LP, Mazzo A, Lopes Neto FD, Boostel R, et al. Validação psicométrica do instrumento Creighton para avaliação de competências clínicas em simulação. Acta Paul Enferm. 2020;33:eAPE20200314.

## DOI

<http://dx.doi.org/10.37689/actaape/2020A003146>



## Descritores

Competência clínica; Psicométrica; Simulação; Estudos de validação

## Keywords

Clinical competence; Psychometrics; Simulation; Validation studies

## Descriptores

Competencia clínica; Psicométrica; Simulación; Estudios de validación

## Submetido

11 de Dezembro de 2019

## Aceito

27 de Abril de 2020

## Autor correspondente

Jéssica de Oliveira Veloso Vilarinho  
E-mail: [jessica.o.veloso@gmail.com](mailto:jessica.o.veloso@gmail.com)

## Resumo

**Objetivo:** Validar construto e avaliar a confiabilidade do Instrumento Creighton para Avaliação de Competências Clínicas, versão português, Brasil.

**Métodos:** Estudo metodológico, realizado com 30 docentes/enfermeiros que utilizavam a metodologia de simulação. Foram elaborados dois cenários com temas iguais e atuações divergentes das estudantes com o objetivo de testar a capacidade do instrumento de diferenciá-los. Os cenários foram gravados e disponibilizados para os especialistas por um *website* que viabilizou a coleta. Para a validação de construto, utilizou-se as técnicas de experimentação e da correlação phi. Para a confiabilidade, analisou-se a equivalência, com coeficiente Fleiss Kappa, e a consistência interna, com coeficiente Kuder-Richardson (KR20).

**Resultados:** Ao analisar a validade, 20 dos 23 itens apresentaram resultados significativos. Na análise, os domínios apresentaram correlações, classificadas de baixa a alta. Quando comparados os cenários, as correlações foram baixas, muito baixas ou negativas. Quanto à confiabilidade, o coeficiente Fleiss Kappa encontrado foi razoável (0,282) para o cenário 1 e moderado para o cenário 2 (0,408). A consistência interna medida com o KR20 foi de 0,717 para o cenário 1 e 0,805 para o cenário 2, classificadas como moderada e alta.

**Conclusão:** O Instrumento Creighton para Avaliação de Competências Clínicas versão português, Brasil, foi considerado válido e confiável para uso em todo o território nacional.

## Abstract

**Objective:** To validate and assess the Creighton Competency Evaluation Instrument reliability, Brazilian Portuguese version.

**Methods:** This is a methodological study conducted with 30 professors/nurses who used the simulation methodology. Two scenarios were elaborated with equal themes and divergent performances of students in order to test the ability of the instrument to differentiate them. The scenarios were recorded and made available to the experts by a website that enabled the collection. For construct validation, the experimentation and phi correlation techniques were used. For reliability, equivalence with the Fleiss Kappa coefficient and internal consistency with Kuder-Richardson coefficient (KR20) were analyzed.

**Results:** When analyzing validity, 20 of the 23 items showed significant results. In analysis, the domains presented correlations, classified from slight to substantial. When the scenarios were compared, the correlations were low, poor or negative. Regarding reliability, the Fleiss Kappa coefficient found was fair (0.282) for scenario 1 and moderate for scenario 2 (0.408). The internal consistency measured with KR20 was 0.717 for scenario 1 and 0.805 for scenario 2, classified as moderate and substantial.

<sup>1</sup>Universidade Federal do Paraná, Curitiba, PR, Brasil.

<sup>2</sup>Universidade de São Paulo, Bauru, SP, Brasil.

Conflitos de interesse: nada a declarar.

**Conclusion:** The Creighton Competency Evaluation Instrument, Brazilian Portuguese version, was considered valid and reliable for use nationally.

## Resumen

**Objetivo:** Validar el constructo y analizar la confiabilidad del instrumento Creighton para Evaluación de Competencias Clínicas, versión en portugués, Brasil.

**Métodos:** Estudio metodológico, realizado con 30 docentes/enfermeros que utilizaban la metodología de simulacro. Se elaboraron dos situaciones con temas iguales y actuaciones divergentes de las estudiantes, con el objetivo de probar la capacidad del instrumento de diferenciarlas. Las situaciones fueron grabadas y se pusieron a disposición de los especialistas mediante un sitio web que viabilizó la recolección. Para la validación del constructo se utilizaron las técnicas de experimentación y del coeficiente de correlación phi. Para la confiabilidad, se analizó la equivalencia con coeficiente Fleiss Kappa y la consistencia interna con coeficiente Kuder-Richardson (KR20).

**Resultados:** Al analizar la validez, 20 de los 23 ítems presentaron resultados significativos. En el análisis, los dominios presentaron correlaciones, clasificadas de baja a alta. Al comparar las situaciones, las correlaciones fueron bajas, muy bajas o negativas. Respecto a la confiabilidad, el coeficiente Fleiss Kappa encontrado fue razonable (0,282) en la situación 1 y moderado en la situación 2 (0,408). La consistencia interna medida con el KR20 fue de 0,717 en la situación 1 y 0,805 en la situación 2, clasificada como moderada y alta.

**Conclusión:** El instrumento Creighton para Evaluación de Competencias Clínicas, versión en portugués, Brasil, fue considerado válido y confiable para su uso en todo el territorio nacional.

## Introdução

A formação de enfermeiros baseada em competências tem sido o foco das reestruturações curriculares desde a incorporação do termo nas Diretrizes Curriculares Nacionais de Enfermagem, de 2001.<sup>(1,2)</sup> A partir de então, houve uma consequente exigência na inclusão de novas práticas pedagógica, com inevitável impacto em avaliações docentes mais rigorosas nas atividades práticas dos estudantes de enfermagem.<sup>(2)</sup>

Nesse contexto de busca por uma metodologia que possibilite a formação por competência, a simulação clínica surge como uma técnica capaz de potencializar a aprendizagem e desenvolver competências clínicas essenciais.<sup>(3,4)</sup> Por *competência clínica*, entende-se como a interligação de conhecimentos, habilidades e atitudes discretas e mensuráveis, essenciais à qualidade do atendimento e à segurança do paciente.<sup>(5)</sup>

A simulação clínica consiste na replicação fidedigna de ambientes clínicos, que permite ao estudante o desenvolvimento de competências em ambiente seguro, com autoconfiança, satisfação e reflexão crítica.<sup>(6)</sup> Ao mesmo tempo, esta estratégia possibilita ao docente uma avaliação com base em critérios pré-estabelecidos, em ambiente controlado, e que pode ser repetido quantas vezes for necessário.<sup>(7)</sup>

No que diz respeito ao processo avaliativo, é importante que as decisões do docente, relacionadas principalmente à avaliação de competências clínicas, sejam orientadas por instrumentos válidos e confiáveis para uso em ambiente simulado.<sup>(8,9)</sup> Os instrumentos são imprescindíveis para o reconhecimento

das lacunas de conhecimento de cada estudante, para uma observação sistematizada e criteriosa das ações e para neutralidade das decisões dos docentes.<sup>(10)</sup>

Nesse sentido, recentemente o instrumento *Creighton Competency Evaluation Instrument (CCEI)*, construído com o objetivo de avaliar competências clínicas em ambiente de simulação, passou pelas etapas de tradução e adaptação transcultural para a língua portuguesa, Brasil, tendo sido intitulado nesta versão como *Instrumento Creighton para Avaliação de Competências Clínicas (CCEI-Br)*, versão português, Brasil.

A avaliação de competências em ambiente simulado ainda é uma prática recente no Brasil, pois há o predomínio do uso da simulação para avaliação de habilidades, com uso de *check-lists* criados pelos próprios docentes, e as competências avaliadas em prática clínica real. Além disso, as pesquisas brasileiros preocupam-se, em sua maioria, em comprovar a eficiência da prática simulada. Esta tradução e adaptação foi importante, portanto, porque até então não havia disponível, no território brasileiro, instrumentos com este objetivo, direcionados para esta prática.<sup>(11,12)</sup>

Assim, o presente trabalho tem como objetivo verificar a validade de construto e avaliar as propriedades de confiabilidade do *Instrumento Creighton para Avaliação de Competências Clínicas (CCEI-Br)*.

## Métodos

Estudo metodológico, de abordagem quantitativa, que seguiu as recomendações do referencial metodológico

de Pasquali.<sup>(13)</sup> O estudo foi desenvolvido no laboratório de simulação clínica de uma universidade pública federal, do sul do Brasil e de forma *on line*, com uso de uma plataforma de criação de *sites* gratuitos, no período de março a julho de 2019. A população do estudo consistiu em enfermeiros e/ou docentes de enfermagem, provenientes de todo o território nacional que utilizavam a metodologia da simulação clínica em suas práticas.

O recrutamento aconteceu via e-mail com o convite a *experts* da área da simulação clínica em enfermagem e a indicação destes; por meio da divulgação em grupos de pesquisa e pela busca ativa na plataforma *lattes*, disponível na página eletrônica do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), com uso da guia *assunto* e com os descritores *simulação clínica e enfermagem*.

A amostra foi selecionada, portanto, com uso da técnica de amostragem não probabilística por conveniência. Um total de 221 e-mails foram disparados, com uma conversão de 30 profissionais que concordaram em participar. O instrumento utilizado para a coleta de dados foi o CCEI-Br, versão traduzida e adaptada para o Brasil do instrumento CCEI original, desenvolvido em 2008 pela Universidade de Creighton, no estado norte-americano de Nebraska.<sup>(14)</sup> Após quatro anos da criação da versão original, o CCEI foi revisado para atender aos conceitos e terminologias do *Quality and Safety Education for Nurses* (QSEN), com 23 comportamentos de classificação dicotômica entre *demonstra e não demonstra competência* (0 a 1), além de *não se aplica*, distribuídos nos domínios: *avaliação, comunicação, julgamento clínico e segurança do paciente*.<sup>(15)</sup>

O CCEI foi validado, com 31 *experts* docentes do território norte-americano. Nesta ocasião, encontrou-se um alfa de Cronbach de 0,979 e um percentual de concordância de 79,4% entre os especialistas.<sup>(11)</sup> No Brasil, na etapa de tradução e adaptação transcultural foi encontrado um alfa de Cronbach de 0,897, considerado uma ótima consistência interna e um percentual de concordância de 100% entre os especialistas.<sup>(14)</sup>

A operacionalização desse estudo aconteceu por etapas, com início na elaboração e gravação de dois cenários de simulação clínica com temas iguais e desempenhos divergentes. Essa necessidade surgiu para responder à hipótese criada de que o instrumento seria capaz de discriminar as atuações divergentes de

dois participantes, em cenários produzidos experimentalmente, conforme recomenda a técnica de *experimentação* de Pasquali para validação de construto.

Os dois cenários gravados eram idênticos e foram baseados em uma prática simulada sobre alteração do nível de consciência, causado por hipoglicemia, que contemplavam todos os itens do CCEI e que se diferenciavam apenas pelo desempenho divergente das duas estudantes. Para a elaboração dos cenários foram seguidas as recomendações metodológicas de Jeffries e de Waxman.<sup>(4,16)</sup> O tema escolhido seguiu as últimas diretrizes da Sociedade Brasileira de Diabetes.<sup>(17)</sup>

Os cenários eram compostos por um paciente (ator), um estagiário (suporte ao estudante) e uma enfermeira (estudante que seria avaliado). Ao iniciar o cenário, a estudante se deparava com o paciente desacordado e o estagiário em desespero pois há pouco havia conversado com o paciente. O objetivo do cenário era que a estudante identificasse a causa da não responsividade do paciente, com identificação dos sinais vitais e verificação da glicemia, e procedesse com a intervenção correta, que seria a administração da glicose, conforme prescrição médica e protocolo de hipoglicemia disponíveis.

Após a gravação audiovisual dos cenários, um pré-teste foi aplicado com cinco docentes/enfermeiros convidados que conheciam o método da simulação, para que avaliassem o desempenho das estudantes nos dois cenários. Na ocasião, os participantes do pré-teste assistiram aos dois cenários gravados e avaliaram o desempenho das estudantes com uso do CCEI-Br. Na sequência, com o objetivo de abranger o maior número de participantes, um *website* foi construído, para que fosse possível disponibilizar à população participante o passo a passo de como responder à pesquisa, os cenários gravados, o CCEI para avaliação dos cenários e guias de orientações.

Na etapa de coleta dos dados, os participantes que concordaram com o Termo de Consentimento Livre e Esclarecido (TCLE) acessaram ao link do *website* disponibilizado no e-mail convite e avaliaram o desempenho das estudantes com apoio da planilha de discussão, previamente elaborada, que continham os comportamentos mínimos esperados de cada estudante, para cada item, com o objetivo de se estabelecer um padrão de avaliação. Os dados coletados foram armazenados em planilha do *Microsoft Office Excel* 2016.

Na análise dos dados foram avaliadas a validade de construto e confiabilidade do instrumento. Para a validade de construto foram escolhidas duas técnicas, dentre as diversas formas de se validar. A primeira diz respeito à *experimentação*, descrita por Pasquali como a melhor técnica para *análise por hipótese*, que é a capacidade do instrumento em discriminar grupos-critérios e nesse caso foi necessário a formulação de uma hipótese: o instrumento seria capaz de discriminar as atuações divergentes de dois participantes, em cenários produzidos experimentalmente. Para este caso, foi utilizado análise estatística comparativa, com o *software Statistical Package for the Social Sciences (SPSS)* versão 25.0.

A segunda técnica foi a *análise comportamental do construto*, em que se analisa a correlação entre os itens, com o cálculo do coeficiente de correlação phi ( $\Phi$ ). Este coeficiente é o de escolha para instrumentos dicotômicos, variando de -1 a +1. A interpretação deste coeficiente seguiu a classificação: muito baixa ( $0 < \phi < 0,1$ ), baixa ( $0,1 \leq \phi < 0,3$ ), moderada ( $0,3 \leq \phi < 0,5$ ), alta ( $0,5 \leq \phi < 0,7$ ), muito alta ( $0,7 \leq \phi < 0,9$ ) e perfeita ( $0,9 \leq \phi \leq 1$ ).<sup>(18)</sup> Na análise estatística foi utilizado o *Software Estatístico Livre R*, com uso do pacote *psych*.

Na avaliação da confiabilidade foram analisados dois critérios, conforme referencial. O primeiro se referiu à equivalência do teste, que tem relação com o grau de concordância entre observadores. Nesse contexto, foi o utilizado o coeficiente de concordância Fleiss Kappa, ideal para análises com mais de dois observadores, com apoio do *Software Estatístico Livre R*, pacote *irr*. Os valores de referência deste coeficiente foram: insignificante ( $< 0,00$ ), leve (0,00 a 0,20), razoável (0,21 a 0,40), moderado (0,41 a 0,60), forte (0,61 a 0,80) e quase perfeito (0,81 a 1,00).

O segundo critério de confiabilidade consistiu na análise da consistência interna. Para tal, foi utilizado o coeficiente de Kuder-Richardson (KR20), específico para testes com itens dicotômicos e que pode variar de 0 a 1. A interpretação dos resultados do KR20 seguiu a classificação: baixa ( $KR20 < 0,50$ ), moderada ( $0,50 \leq KR20 < 0,8$ ) e alta ( $KR20 \geq 0,80$ ). Nesse caso, a análise estatística foi realizada com auxílio do *Software Estatístico Livre R*, pacote *DescTools*.

O presente estudo atendeu aos critérios contidos na resolução nº 466, de 12 de dezembro de 2012, com aprovação do estudo no Comitê de Ética

e Pesquisa da Universidade Federal do Paraná, sob o número de parecer 3.175.444.

## Resultados

A caracterização sociodemográfica apresentou-se com maioria do sexo feminino (80%), com vínculos em instituições públicas (76,7%) e com ao menos o mestrado como formação (90%). A experiência de no mínimo dois anos com a simulação foi comprovada por 76,7% dos participantes, dos quais 69,6% possuíam mais de cinco anos. Quase todas as regiões do Brasil possuíam representantes na amostra, com exceção da região norte.

Na etapa de validação do construto, com *análise por hipótese*, 20 dos 23 itens apresentaram divergência entre os dois cenários, com resultados significativos ( $p < 0,05$ ). O domínio *juízo clínico* apresentou resultados significativos em oito dos nove itens, com exceção do item 16 (faz reflexão sobre a experiência clínica) ( $p = 0,456$ ). No domínio *segurança do paciente*, dois dos seis itens apresentaram resultados que rejeitam a hipótese elaborada ( $p > 0,05$ ), dos quais foram: o item 20 (administra medicamentos com segurança) e o item 23 (reflete sobre potenciais erros e riscos).

Quanto à análise comportamental do construto, houve correlação entre todos os itens, que variaram de muito baixa a alta. O domínio *avaliação* demonstrou coeficientes de correlação phi de moderados a alto ( $\phi$ : 0,370 a 0,557), tanto no cenário 1 como no cenário 2. No domínio *comunicação*, o item 8 (atua de forma profissional) do cenário 2 não apresentou correlação, por invariância nas respostas. Os demais itens apresentaram predominância de resultados considerados baixos, com valores de phi entre 0,1 e 0,3, para ambos os cenários.

Para o domínio *juízo clínico*, as correlações variaram de muito baixa a alta, com predomínio da correlação baixa no cenário 1 e da moderada no cenário 2. Os resultados do domínio *segurança do paciente* apontaram correlações similares ao domínio *juízo clínico*. Os dois cenários foram comparados, item a item, quanto ao coeficiente de correlação phi, conforme a tabela 1, com a maioria dos itens apresentando correlação muito baixa ou negativa.

Quanto à equivalência para teste de confiabilidade, foram calculados o grau de concordância Fleiss Kappa e

**Tabela 1.** Matriz de correlação do coeficiente phi comparativa entre os itens dos cenários 1 e 2

Item do Cenário 1 versus do Cenário 2	Coefficiente de correlação phi	Item do Cenário 1 versus do Cenário 2	Coefficiente de correlação phi
1	0,062	13	N/A
2	N/A	14	-0,141
3	0,062	15	0,060
4	0,034	16	-0,449
5	0,064	17	0,208
6	0,204	18	-0,141
7	0,073	19	-0,256
8	N/A	20	-0,254
9	-0,068	21	0,090
10	0,152	22	-0,099
11	-0,082	23	0,156
12	N/A	-	-

N/A – não foi possível calcular a correlação

o percentual de concordância, por domínio e para cada cenário, conforme demonstra a tabela 2 abaixo. Os cenários 1 e 2 apresentaram grau de concordância Fleiss Kappa de 0,282 e 0,408, considerados respectivamente razoável e moderado e percentual de concordância dos cenários 1 e 2 foi 81% e 63%, respectivamente.

Quanto à consistência interna para mensurar a confiabilidade do instrumento CCEI, observa-se na tabela 3 que o coeficiente de fidedignidade KR20 foi de 0,717 para o cenário 1, considerado moderado, e de 0,805 para o cenário 2, classificado como de alta consistência. Todos os domínios, se excluídos, repercutem negativamente no escore geral de fidedignidade da escala.

**Tabela 2.** Grau de concordância Fleiss Kappa e percentual de concordância para cada cenário

		Total de itens	Fleiss kappa	p-value	% concordância	Grau de concordância
CENÁRIO 1	GERAL	23	0,282	<0,001*	81	Razoável
	Avaliação	3	0,022	0,392	89	-
	Comunicação	5	0,356	<0,001*	88	Razoável
	Julgamento clínico	9	0,173	<0,001*	86	Leve
	Segurança do paciente	6	0,281	<0,001*	63	Razoável
CENÁRIO 2	GERAL	23	0,408	<0,001*	63	Moderado
	Avaliação	3	0,008	0,783	96	-
	Comunicação	5	0,546	<0,001*	78	Moderado
	Julgamento clínico	9	0,225	<0,001*	61	Razoável
	Segurança do paciente	6	0,289	<0,001*	36	Razoável

p – Teste de McNemar

**Tabela 3.** Coeficiente de confiabilidade Kuder-Richardson para cada cenário e por domínio

Intervalo de confiança (95%)		Cenário 1 KR20	Cenário 2 KR20
<b>GERAL</b>		<b>0,717</b>	<b>0,805</b>
AVALIAÇÃO	ITENS	0,633	0,494
	1 – Obtem dados pertinentes	0,434	0,000
	2 – Realiza avaliação e acompanhamento conforme necessário	0,658	0,659
	3 – Avalia o ambiente de forma organizada	0,541	0,000
COMUNICAÇÃO	ITENS	0,201	0,404
	4 – Comunica-se efetivamente com equipe intra/interprofissional	0,215	0,340
	5 – Comunica-se efetivamente com o paciente e acompanhante (verbal, não verbal, fornece orientações)	0,137	0,187
	6 – Documento de forma clara, concisa e precisa	-0,058	0,369
	7 – Responde apropriadamente a achados anormais	0,137	0,381
	8 – Atua de forma profissional	0,215	0,431
JULGAMENTO CLÍNICO	ITENS	0,601	0,749
	9 – Interpreta sinais vitais (T, PA, FR, FC, Dor)	0,583	0,734
	10 – Interpreta resultados laboratoriais	0,609	0,725
	11 – Interpreta dados objetivos/subjetivos (distingue entre dados relevantes e irrelevantes)	0,510	0,729
	12 – Prioriza as ações de forma adequada	0,528	0,761
	13 – Executa intervenções baseadas em evidências	0,610	0,661
	14 – Apresenta fundamentação baseada em evidências para as intervenções	0,610	0,691
	15 – Avalia as intervenções baseadas em evidências e seus resultados	0,512	0,742
	16 – Faz reflexão sobre a experiência clínica	0,497	0,751
17 – Delega apropriadamente	0,627	0,718	
SEGURANÇA DO PACIENTE	ITENS	0,275	0,625
	18 – Utiliza identificadores de pacientes	0,364	0,636
	19 – Utiliza práticas padronizadas e precauções, incluindo a higienização das mãos	0,127	0,586
	20 – Administra medicamentos com segurança	-0,104	0,496
	21 – Utiliza tecnologia e equipamentos adequadamente	0,253	0,598
	22 – Executa procedimentos corretamente	0,475	0,602
23 – Reflete sobre potenciais erros e riscos	0,069	0,552	

KR20 – Coeficiente de Kuder-Richardson

A categoria *segurança do paciente* é a que apresenta maior variação no KR20, de 0,275 no cenário 1, para 0,625 no cenário 2. Já a categoria *juízo clínico* é a que apresenta os melhores valores de confiabilidade, se analisado isoladamente, com KR20: 0,601 para o cenário 1 e KR20: 0,749 para o cenário 2.

## Discussão

Os resultados deste estudo, permitiram a validação de construto e a confirmação da confiabilidade do instrumento CCEI, o que permite disponibilizar aos docentes e enfermeiros de todo o território nacional, um instrumento capaz de estimar a competência clínica de estudantes e profissionais de enfermagem que forem avaliados em cenários de simulação clínica de alta fidelidade. Acredita-se que esses resultados poderão contribuir especialmente para o desenvolvimento de competências clínicas ainda na formação dos estudantes de enfermagem.

No que se refere à caracterização sociodemográfica, a amostra consistiu em sua maioria em participantes do sexo feminino, característica muito relacionada à profissão de enfermagem até os dias de hoje, apesar da proporção do sexo masculino ter aumentado em salas de aulas dos cursos de graduação.<sup>(19)</sup> Os docentes e enfermeiros participantes eram provenientes, prevalentemente, de instituições públicas e possuíam mestrado como último grau de instrução. Esse resultado é justificado com a predominância da produção científica nas instituições públicas.<sup>(20)</sup>

Na validação de construto, levou-se em consideração as análises conceituais de cada domínio. Em relação à análise por hipótese, os domínios *avaliação e comunicação* apresentaram resultados que confirmaram a hipótese elaborada. Relacionado à análise comportamental do construto, estes mesmos domínios apresentaram coeficientes de correlação phi moderados e altos para *avaliação* e baixos para *comunicação*, em ambos os cenários. Quando comparados os mesmos itens destes domínios nos dois cenários, obteve-se uma correlação muito baixa, o que permite concluir que os domínios supracitados comprovaram a validação de construto.

O coeficiente de correlação phi é considerado o de escolha para a teoria psicométrica mais atual, denominada Teoria de Resposta ao Item, e é utilizado, com frequência, em análises fatoriais, que não foram analisadas neste estudo. Por esse motivo, é necessário salientar que, em geral, são recomendadas amostras superiores a 50 participantes, o que não foi o caso e que pode justificar os resultados abaixo do esperado para este coeficiente.<sup>(21)</sup> Além disso, a quantidade de itens em cada domínio influencia diretamente nos resultados do coeficiente de correlação.<sup>(13)</sup> Diante desse contexto, ressalta-se que o coeficiente de correlação phi não foi verificado isoladamente neste estudo.

No que se refere ao domínio *juízo clínico*, o item 16 (faz reflexão sobre a experiência clínica) foi o único que apresentou resultados não significativos, devido à quantidade de opções *não se aplica* assinaladas. Isso aconteceu porque os avaliadores, por dominarem o método da simulação, relataram na *observação* que este item deveria ser analisado no momento do *debriefing*, que é a etapa que acontece após o cenário simulado, e que não foi abordada na gravação.

Quanto ao coeficiente de correlação do domínio *juízo clínico*, houve prevalência de correlações baixas a moderadas em ambos os cenários, e muito baixo ou negativo quando tem-se os cenários comparados. Nota-se que consistências internas mais fortes foram ressaltadas nos resultados do cenário 2, relacionados aos coeficientes de correlação, justificadas pelo julgamento facilitado do mau desempenho da estudante nesse cenário, o que resultou em mais respostas: *não demonstra competência* e menos *não se aplica*.

Sobre a avaliação do cenário 1, relacionado a este domínio, percebeu-se que apesar das orientações de seguir os comportamentos mínimos, previamente definidos, os enfermeiros inevitavelmente julgaram sob a perspectiva de suas formações e valores e encontraram dificuldades em estabelecer padrão de respostas, conforme confirma a literatura.<sup>(22)</sup> Com base nisso, pode-se concluir que a categoria da competência *juízo clínico* demonstrou achados que permitem confirmar a validação de construto.

Sobre o domínio *segurança do paciente*, somente os itens 20 (administra medicamentos com segurança) e 23 (reflete sobre potenciais erros e riscos) não apresen-

taram resultados significativos, relacionados à hipótese de divergência. O item 23 apresenta a mesma justificativa do item 16 e está relacionado ao *debriefing*. Com relação ao item 20 (administra medicamentos com segurança), houve divergências de opiniões sobre o desempenho da estudante do cenário 1.

Os participantes relataram no campo *observação* do instrumento: “senti dificuldades em responder alguns quesitos, pois abrangiam mais que uma afirmativa e às vezes apenas uma tinha sido seguida corretamente. Por exemplo, ela não realiza a lavagem das mãos, mas utiliza a luva” (participante 13). O tema segurança do paciente consiste em complexa discussão, que envolve uma das metas internacionais para garantia da segurança do paciente, instituída pela Organização Mundial da Saúde, e portanto, passível de maior rigor na avaliação de desempenhos.<sup>(23)</sup>

Com relação ao coeficiente de correlação deste domínio, encontraram-se coeficientes mais fortes no cenário 2, assim como nos demais domínios e prevalência de correlações baixas a moderadas. Quando comparados os cenários, os itens deste domínio apresentaram correlações muito baixas ou negativas, o que permite concluir que *segurança do paciente* apresentou resultados que confirmam a validação de construto.

A confiabilidade do instrumento CCEI, relacionada a equivalência, foi confirmada com o coeficiente de concordância Fleiss Kappa considerados razoável para o cenário 1 e moderado para o cenário 2, com valores de 0,282 e 0,408. Na validação da escala original, os autores encontraram coeficientes de concordância classificados também como razoáveis e moderados, com valores 0,316, 0,453 e 0,443.<sup>(15)</sup>

O CCEI validado na versão espanhola encontrou valores de coeficiente de concordância Kappa de 0,80, porém ressalta-se que o estudo finalizou a escala com apenas 22 itens e o método de análise de escolha foi diferente do estudo atual.<sup>(24)</sup> Os coeficientes Fleiss Kappa dos domínios desse estudo apresentaram resultados inversamente proporcionais com os percentuais de concordância, justificados pela análise de existência de erro aleatório que o coeficiente Fleiss Kappa utiliza, ao contrário do percentual de concordância, que apenas avalia a concordância bruta dos avaliadores.<sup>(25)</sup>

A consistência interna, calculada com o coeficiente de confiabilidade KR20, apresentou valores de KR20:0,717 para o cenário 1 e KR20: 0,805 para o cenário 2, classificadas como consistência interna moderada e alta, respectivamente. A versão original do CCEI encontrou valores de consistência interna maiores que 0,90. Ressalta-se que o coeficiente utilizado neste caso foi o Alfa de Cronbach.<sup>(15)</sup>

Na versão espanhola, a consistência interna encontrada foi de 0,839, classificada como alta.<sup>(24)</sup> Sobre a versão brasileira, ainda na etapa de tradução, foi encontrado Alfa de Cronbach de 0,897.<sup>(14)</sup> Alguns autores encontraram valores menores de consistência interna, quando utilizado o coeficiente KR20, com valores de 0,701 e 0,803, para escalas dicotômicas, considerado consistência moderada e alta, respectivamente.<sup>(26)</sup>

O processo de validação e de verificação da confiabilidade seguiu as orientações do referencial metodológico e apresentou limitações no que se refere à disponibilidade de outros instrumentos válidos e fidedignos, construídos com o mesmo objetivo, e que possibilitassem a verificação de outros tipos de validação. Além disso, o estudo apresentou limitações quanto ao recrutamento dos participantes, que por ter sido via e-mail, acabou permitindo uma baixa conversão dos que aceitaram contribuir com a pesquisa.

## Conclusão

O presente estudo atendeu ao objetivo de validar o construto e testar a confiabilidade da versão em português, Brasil, do instrumento CCEI. As análises estatísticas e teóricas realizadas comprovam que o instrumento foi capaz de discriminar os cenários divergentes. Além disso, os achados dos coeficientes de correlação foram considerados aceitáveis, em associação à discussão teórica. Estes resultados confirmaram a validação de construto do instrumento. A análise da confiabilidade demonstrou parâmetros de equivalência considerados, em geral, razoável e moderado, para os dois cenários e parâmetros de consistência interna considerados moderados e altos. Dessa forma, pode-se concluir que o referido instrumento apresenta boa consistência interna e

é considerado confiável para medir a competência clínica de estudantes/profissionais de enfermagem, quando avaliados em simulação clínica.

## Colaborações

Vilarinho JOV, Felix JVC, Kalinke LP, Mazzo A, Lopes Neto FDN, Boostel R, Silva NO e Fontoura ACOB declaram que contribuíram com a concepção do projeto, análise e interpretação dos dados, redação do artigo, revisão crítica relevante do conteúdo intelectual e aprovação da versão final a ser publicada.

## Referências

1. Brasil. Conselho Nacional da Educação. Câmara de Educação Superior. Resolução CNE/CES 3/2001. Brasília (DF): Ministério da Saúde; 2001. [citado 2019 nov 01]. Disponível em: <http://portal.mec.gov.br/cne/arquivos/pdf/CES03.pdf>
2. Winters JR, Padro ML, Lazzari DD, Jardim VL. A formação superior em enfermagem no MERCOSUL: um estudo bibliométrico [Internet]. Rev Bras Enferm. 2018;71 Suppl 4:1732–9.
3. Rosa RS, Sanches GJC, Gomes ICR, Silva ML, Duarte ACS, Boery RNSO. Estratégias baseadas em metodologias ativas no ensino-aprendizagem de primeiros socorros: relato de experiência. Rev Enferm UFPE Online. 2017;11(2):798-803.
4. Jeffries PR, Rodgers B, Adamson K. NLN Jeffries Simulation Theory: brief narrative description. Nurs Educ Perspect. 2015 ;36(5):292–3.
5. Meakim C, Boese T, Decker S, Franklin AE, Gloe D, Lioce L, et al. Standards of best practice: simulation. Clin Simul Nurs. 2013;9(6):3–11.
6. Costa RRO, Medeiros SM, Martins JCA, Enders BC, Lira ALBC, Araújo MS. A simulação no ensino de enfermagem: uma análise conceitual. Rev Enferm Centro-Oeste Mineiro. 2018;8(e1928):e19288.
7. Escudero E, Bem-Azul MA, Cancino KD. Simulación clínica y seguridad del paciente: integración em el currículo de enfermería. Sci Med (Phila). 2018;28(1):1–13.
8. Bergamasco EC, Murakami BM, Cruz DA. Uso da Escala de Satisfação dos Estudantes e Autoconfiança com a Aprendizagem (ESEAA) e da Escala do *Design* da Simulação (EDS) no ensino de enfermagem: relato de experiência. Sci Med (Phila). 2018;28(3):1–5.
9. Presado MH, Colaço S, Rafael H, Baixinho CL, Félix I, Saraiva C, et al. Aprender com a simulação de alta fidelidade. Cien Saude Colet. 2018;23(1):51–9.
10. Gomes JÁ, Martins MM, Tronchin DM, Fernandes CS. Validação de uma escala sobre indicadores de estrutura no bloco operatório: contributos para a enfermagem. Arquichan. 2019;19(1):1–13.
11. Miranda FBG, Mazzo A, Pereira Junior GA. Avaliação de competências individuais e interprofissionais de saúde em atividades clínicas simuladas: scoping review. Interface comum. Saúde Educ. 2018;22(67):1221-34.
12. Sebold LF, Boell JEW, Girondi JBR, Santos JLG. Simulação clínica: desenvolvimento de competência relacional e habilidade prática em fundamentos de enfermagem. Rev Enferm UFPE Online. 2017;11(Supl 10):4184-90.
13. Pasquali L. Validade dos testes. Revista Examen. 2017;1(1):14–48.
14. Silva NO. Tradução e adaptação transcultural do Creighton Competency Evaluation Instrument para a língua portuguesa (Brasil) [dissertação]. Curitiba: Universidade Federal do Paraná; 2019.
15. Hayden J, Keegan M, Kardong-Edgren S, Smiley RA. Reliability and validity testing of the Creighton Competency Evaluation Instrument for use in the NCSBN National Simulation Study. Nurs Educ Perspect. 2014;35(4):244–52.
16. Waxman KT. The development of evidence-based clinical simulation scenarios: guidelines for nurse educators. J Nurs Educ. 2010;49(1):29–35.
17. Oliveira JP, Motenegro RM Junior, Vencio S. Diretrizes da Sociedade Brasileira de Diabetes 2017-2018. São Paulo: Editora Clannad; 2017.
18. Burns MK, Zaslofsky AF, Kanive R, Parker DC. Meta-analysis of incremental rehearsal using phi coefficients to compare single-case and group designs. J Behav Educ. 2012;21(1):158–202.
19. Correa Júnior AJ, Santana ME, Farias GM, Noronha EC. Globalización em la postmodernidad: críticas y contribuciones para la enfermería. Cultura de los Cuidados. 2019;23(53):265-74.
20. Smolski FMS, Dalcin D, Visentini MS, Bamberg J, Kern JS. Análise do perfil da produção científica da Revista de Administração Pública (RAP). Rev. adm. pública. 2020. [cited 2020 abr. 15]. Available from: [http://www.scielo.br/pdf/rap/v51n6/1982-3134-rap-51-06-1139.pdf](http://www.scielo.br/pdf/rap/v51n6/1982-3134-1982-3134-rap-51-06-1139.pdf)
21. Collares CF, Grec WL, Machado JL. Psicometria na garantia de qualidade da educação médica: conceitos e aplicações. Science in Health. 2012;3(1):33–49.
22. Santos VR. Avaliação escolar – uma prática entre duas lógicas: a que emancipa e a que reproduz as desigualdades sociais. Interação (Curitiba). 2019;44(2):515–28.
23. Cardoso AS, Muller S, Echer IC, Rabelo-Silva ER, Boni FG, Ribeiro AS. Elaboração e validação do *checklist* para administração de medicamentos para pacientes em protocolo de pesquisa. Rev Gaúcha Enferm. 2019;40(spe):1-5.
24. Roldán-Merino J, Farrés-Tarafa M, Estrada-Masllorens JM, Hurtado-Pardos B, Miguel-Ruiz D, Nebot-Bergua C, et al. Reliability and validity study of the Spanish adaptation of the “Creighton Simulation Evaluation Instrument (C-SEI)”. Nurse Educ Pract. 2019;35(1):14–20.
25. Kuo BI. Intraclass correlation coefficient rather than correlation coefficient to examine agreements among different methods measuring valvular area. Circulation. 1994;89(4):1910–1.
26. Vizzotto PA, Mackedanz LF. Validação de instrumento de avaliação da alfabetização científica para egressos no ensino médio no contexto da física do trânsito. Educ Rev. 2018;34:e202974.