

VALIDAÇÃO DE MÉTRICAS PARA AVALIAR A QUALIDADE EM INSTITUIÇÕES PÚBLICAS DE ENSINO SUPERIOR

Dagmar Willamowius Vituri*
Daisy Maria Rizatto Tronchin**

RESUMO

Introdução: No campo da educação superior pública, a complexidade em avaliar o alcance dos objetivos institucionais e as incertezas acerca das métricas adotadas pelo Estado prejudicam a implementação de processos avaliativos. **Objetivo:** Construir e validar o conteúdo de indicadores de qualidade para avaliação de Instituições Públicas de Ensino Superior (IES). **Método:** Estudo metodológico desenvolvido em seis etapas, com a participação de 10 juízes de IES, entre 2015 e 2018. Empregou-se a Técnica Delphi e o cálculo do *Content Validity Ratio*, com valor maior ou igual a 0,62. Foram propostos 46 indicadores nas subdimensões: Ensino (16); Pesquisa, extensão, conhecimento e inovação (15); e Gestão (15). **Resultados:** Constata-se a evidência de validade de conteúdo de 38 indicadores, dos quais 11 de estrutura, 10 de processos e 17 de resultados, sendo a subdimensão Gestão a que obteve menores valores de CVR. **Considerações Finais:** Obtiveram-se evidências de validade de uma matriz de 38 indicadores quantitativos, formulados como equações matemáticas, que podem ser utilizados sistematicamente na avaliação interna institucional, complementando a avaliação qualitativa.

Palavras-chave: Organização e administração. Estudos de Avaliação como Assunto. Indicadores. Estudos de validação. Educação superior.

INTRODUÇÃO

A preocupação com a qualidade e com o desenvolvimento de uma cultura avaliativa institucional, na atualidade, faz parte do cotidiano da prática de organizações do ramo da indústria, comércio e prestação de serviços.

No campo da educação superior (ES), a avaliação da qualidade e desempenho institucional tornou-se uma manifestação de responsabilidade política, que favorece a regulação em prol dos interesses da sociedade⁽¹⁾. Ganhou vulto e passou a ser apresentada como uma prioridade por parte das autoridades educacionais, educadores e gestores na década de 1990, com a expansão do acesso à educação superior^(2,3), sobretudo, em razão do aumento do número de instituições privadas⁽³⁾.

A avaliação da ES, além de um procedimento metodológico, é também uma atividade estratégica e política, portanto, se faz necessário considerar aspectos teóricos e éticos⁽⁴⁾. Visa à melhoria do ensino, pesquisa, extensão e gestão e, além disso, vem sendo entendida como uma exigência da sociedade, tendo em vista orientar sua escolha por instituições bem conceituadas e posicionadas em rankings nacionais e internacionais⁽⁵⁾.

Além do enfoque de monitoramento e avaliação, há de se destacar o caráter formativo dos processos

avaliativos, que propiciam autorreflexão, autoconhecimento e aprendizado, a partir do estabelecimento da relação entre processos e resultados⁽⁶⁾. Contudo, avaliar a qualidade da educação superior é uma tarefa complexa frente à multidimensionalidade do construto qualidade, assim como da própria universidade, uma vez que são várias as funções e atividades que a representam, como o ensino, programas acadêmicos, pesquisa e fomento da ciência, extensão e ambiente acadêmico em geral.

Durante o percurso histórico das políticas avaliativas no país, de um processo fragmentado evoluiu-se para um sistema de avaliação intitulado Sistema Nacional de Avaliação da Educação Superior - SINAES, pautado por princípios e dimensões e fundamentado na avaliação institucional interna e externa, além da avaliação dos cursos e do desempenho acadêmico dos estudantes pelo Exame Nacional de Desempenho dos Estudantes - ENADE^(7,8).

O processo é meticuloso e desenvolvido em várias etapas. Quanto aos resultados, além de subsidiarem os órgãos reguladores do Ministério da Educação e Cultura - MEC e agências de fomento, preconiza-se analisá-los internamente à universidade, na identificação de oportunidades de melhoria.

*Enfermeira. Doutora. Hospital Universitário da Universidade Estadual de Londrina. Londrina, PR, BR. E-mail: dagmar@uel.br ORCID ID: 0000-0002-7663-1736.

**Enfermeira. Livre Docente. Professora Associada da Escola de Enfermagem da Faculdade de Medicina de São Paulo. Cidade, SP, BR. E-mail: daisyrt@usp.br ORCID ID: 0000-0003-3192-1956.

Os procedimentos avaliativos do SINAES estão estabelecidos segundo critérios e indicadores formulados como assertivas, para as quais o avaliador deve atribuir um conceito, em ordem crescente de excelência, a partir de uma escala *Likert* de cinco pontos. Contudo, percebe-se que indicadores representados por equações, cujos resultados são expressos em índices e taxas, proporcionam uma interpretação mais objetiva dos resultados, otimizando o processo para a tomada de decisão e o monitoramento contínuo⁽⁹⁾.

Assim, o desafio desta pesquisa está em propor uma matriz de indicadores, na forma de equações matemáticas, a partir dos critérios adotados nos modelos brasileiros de avaliação de instituições de ensino superior, avaliação de cursos e acreditação internacional de cursos, que permita inteligibilidade e comparabilidade e que promova o autoconhecimento relacionado a aspectos de estrutura, processos e resultados.

Em razão da existência de distinção entre as organizações de ensino públicas e privadas, no que tange aos valores e objetivos políticos, sociais e educacionais⁽¹⁰⁾, a aplicabilidade da matriz dos indicadores apresentados para os procedimentos de validação na presente pesquisa está direcionada às instituições públicas de ensino superior.

Nesse sentido, delineou-se como questão de estudo: a matriz de indicadores de qualidade proposta permite monitorar e avaliar o desempenho de componentes de estrutura, processos e resultados em instituições públicas de ensino superior no contexto nacional?

Frete à importância da disseminação da cultura avaliativa proativa e formativa, é que se presume a realização desta pesquisa, que tem por objetivo: desenvolver e validar o conteúdo de indicadores para avaliação da dimensão desempenho do construto qualidade em instituições públicas de ensino superior brasileiras, relacionada a componentes da estrutura, processos e resultados, a partir dos modelos de avaliação adotados pelo Estado.

MÉTODO

Trata-se de uma pesquisa metodológica que visa à elaboração e validação de instrumentos com vistas à solução de problemas concretos do cotidiano⁽¹¹⁾.

O referencial metodológico adotado abarcou seis etapas⁽¹²⁾, a saber:

I - Estabelecimento da estrutura conceitual - delimitação para fins da presente pesquisa da dimensão desempenho, do construto qualidade, o qual é operacionalizado por meio de variáveis, indicadores e índices, que correspondem à maneira como a organização se comporta, tendo em vista sua eficiência e rendimento⁽¹³⁾.

II- Definição dos objetivos do instrumento - avaliação de aspectos da estrutura, processos e resultados, oriundos do modelo avaliativo Donabedian para instituições públicas de ensino superior, na dimensão desempenho do construto qualidade, nas subdimensões: 1-Ensino, 2-Pesquisa, extensão, conhecimento e inovação e 3-Gestão⁽¹⁴⁾.

III- Construção dos itens para cada indicador e das escalas de resposta - o desenvolvimento dos indicadores ocorreu a partir de revisão da literatura pertinente, subsidiado pelos modelos de avaliação de instituições de ensino superior (IES), avaliação de cursos, acreditação internacional de cursos, das diretrizes do Plano Nacional de Educação 2010-2020, do Plano Plurianual 2012-2015 e da experiência de docentes e técnicos de uma universidade pública do estado do Paraná com expertise em avaliação institucional. Essa fase abarcou o período de novembro de 2015 a setembro de 2016 e compreendeu o desenvolvimento, análise e julgamento dos 103 indicadores formulados como equações matemáticas e suas fichas técnicas, organizadas da seguinte maneira para cada indicador: descritor, fundamentação, equação para o cálculo, fonte de dados, critérios de avaliação e amostra, avaliando-se o seu conteúdo e atributos. Participaram seis especialistas (docentes e técnicos), exercendo funções de Pró-reitoria, diretoria e gerência.

IV - Seleção e organização dos itens - considerando-se o percentual de concordância ($\geq 80\%$)⁽¹²⁾ e de posse das contribuições dos juízes, houve a exclusão de 60 indicadores, tendo em vista não terem sido considerados pertinentes para a avaliação da dimensão desempenho do construto qualidade de instituições de educação superior pública, o que vem ao encontro das recomendações de limitar o número métricas tendo em vista não comprometer a operacionalização e perenidade do procedimento avaliativo⁽¹⁵⁾,

V -Estruturação do instrumento – nessa fase, compreendida entre outubro de 2016 a dezembro de 2017, os 46 indicadores foram reestruturados nas Subdimensões⁽¹⁴⁾ Ensino -16 indicadores, Pesquisa,

extensão, conhecimento e inovação - 15 indicadores e Gestão - 15 indicadores, sendo 14 indicadores de estrutura, 15 de processo e 17 de resultados, conforme o modelo Donabedian de avaliação (Figura 1).

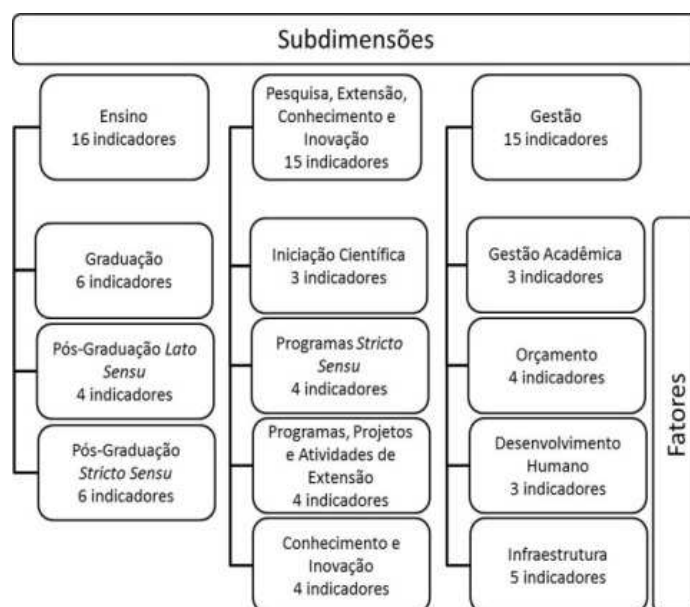


Figura 1. Representação das Subdimensões, fatores e número de indicadores propostos para avaliação da qualidade de instituições públicas de ensino superior, São Paulo – SP, 2015-2016.

Fonte: Próprios autores.

VI-Validade de Conteúdo - procedeu-se à validação de conteúdo por *experts* em ensino, gestão e avaliação da ES, por meio da Técnica Delphi, que tem como base o anonimato, sendo empregada para a obtenção de consenso entre especialistas em determinado campo do conhecimento⁽¹⁶⁾.

A validade de conteúdo alude ao quanto uma amostra de itens é representativa do construto de interesse e, para tanto, depende do julgamento de especialistas, pois não existe teste estatístico para determinar se a medida abarca o conteúdo ou representa adequadamente o construto de interesse⁽¹⁷⁾. Os atributos empregados para avaliação de cada indicador foram: concorrente, objetivo, relevante, comunicável, exequível e acessível; e as respostas de cada especialista ficaram registradas a partir de uma escala dicotômica Sim e Não.

Nesta etapa, o universo amostral foi composto por docentes de instituições públicas de ES do Brasil, atuantes nas áreas acadêmicas ou gestão, especialistas em avaliação do ensino superior e que compunham as Comissões Permanentes de Avaliação, ou equivalente, em sua instituição. A amostragem foi intencional, em razão de selecionar

indivíduos conhecedores das questões que estão sendo estudadas⁽¹¹⁾.

A relação de juízes foi obtida a partir da lista de universidades públicas brasileiras, disponível no Cadastro e-MEC de Instituições e Cursos de Educação Superior, no site do MEC (<http://emec.mec.gov.br/>). Foi gerada uma relação de 107 universidades públicas e, para selecionar as instituições que participariam do estudo, empregaram-se estes critérios de inclusão: todos os estados com menos de quatro universidades públicas; estados acima desse número, considerar: os quatro pontos cardeais principais (norte, sul, leste e oeste), incluindo capital e interior, ao menos uma instituição Federal, uma Estadual e uma Municipal.

Assim, foram selecionadas 74 instituições, contatadas via telefone, solicitando-se o contato do responsável pela Comissão Permanente de Avaliação, para os quais foram apresentados os objetivos e a metodologia da pesquisa e o convite de participação. Aos que aceitaram foi encaminhado, por email, o *link* para acesso ao formulário eletrônico e o Termo de Consentimento Livre Esclarecido, sendo estabelecido o prazo de 30 dias

para o retorno. Houve prorrogação do prazo, tendo em vista o retorno de apenas seis formulários.

O projeto de pesquisa foi submetido ao Comitê de Ética em Pesquisa e aprovado - CAAE nº 57962416.7.0000.5231.

Para a análise de dados, foi realizado o teste da força científica dos indicadores, ou análise teórica dos itens, pelo cálculo *Content Validity Ratio* (CRV), que é uma técnica estatística para determinar se os indicadores são úteis ou não, passíveis ou não de permanecerem no instrumento avaliativo^(18,19).

O cálculo do CVR é obtido com base na resposta do painel de juízes, a partir da fórmula na qual Ne corresponde ao número de juízes que consideraram o item essencial e N o número de juízes que avaliaram o item⁽¹⁸⁾:

$$CRV = \frac{Ne - (N/2)}{N/2}$$

O CVR crítico definido para o presente estudo foi 0,620, com base no tamanho da amostra de juízes^(18,19). Também foi estabelecido que

indicadores com $CVR \geq 0,620$ seriam mantidos, com CVR entre 0,300 e 0,620 reformulados a partir dos comentários e sugestões dos especialistas e com $CVR \leq 0,300$ seriam excluídos do conjunto de indicadores.

RESULTADOS

Responderam o questionário de forma completa 10 especialistas (13,5%). Quanto à organização administrativa, 10% das universidades representadas eram municipais, 40% estaduais e 50% federais, sendo da região norte do Brasil 10%, nordeste 40%, centro oeste 10%, sudeste 30% e sul 10%. Em relação ao perfil dos juízes, 60% eram do sexo masculino, 50% com faixa etária acima de 50 anos, 50% atuavam tanto na gestão como na academia e 50% possuíam tempo de atuação profissional acima de 21 anos.

As Tabelas 1, 2 e 3 apresentam os valores de CRV dos indicadores propostos.

Tabela 1. Distribuição dos indicadores conforme o valor de CRV, na subdimensão Ensino, São Paulo-SP, 2018.

Indicadores		CVR
Subdimensão Ensino		
Fator Graduação		
1	Preenchimento de vagas para ingresso	0.714
2	Número de estudantes por docente	0.714
3	Número de estudantes por docente em regime de tempo integral	0.543
4	Taxa de evasão total	0.800
5	Índice diplomados evadidos	0.457
6	Taxa de conclusão de curso	0.857
Fator Pós-graduação <i>lato sensu</i>		
7	Preenchimento de vagas para ingresso	0.743
8	Taxa de evasão total	0.771
9	Índice titulados evadidos	0.600
10	Taxa de conclusão de curso	0.829
Fator Pós-graduação <i>stricto sensu</i>		
11	Taxa de preenchimento de vagas para ingresso	0.686
12	Taxa de evasão total	0.771
13	Índice titulados evadidos	0.457
14	Taxa de conclusão de curso	0.886
15	Taxa de docentes em processo de internacionalização	0.257
16	Taxa de pós-graduandos em processo de internacionalização	0.343

Na Subdimensão Ensino, no Fator Graduação, o Indicador 3 obteve CVR de 0,543 e, a partir das sugestões dos juízes (J5 e J6), o índice de conformidade foi alterado para “Comparação com o desempenho das IES, no Censo da Educação Superior, bem como com a série histórica dos últimos três anos”.

Quanto aos indicadores 5, 9 e 13 dos fatores Graduação, Pós-Graduação *lato sensu* e *stricto sensu*, os valores de CVR, respectivamente, foram 0,457, 0,600 e 0,457. Os comentários dos especialistas (J5, J6, J8, J9) referiram-se à imprecisão do indicador, tendo-se em vista a série de fatores relacionados à evasão de cursos de

graduação e pós-graduação.

Nesse sentido, este indicador foi reformulado para possibilitar mensurar separadamente os diferentes atributos que compõem o construto evasão, a saber: cancelamento, desistência, trancamento total, transferência externa e jubramento.

Quanto aos indicadores 15 e 16, o CVR obtido foi 0,257 e 0,043, respectivamente, o que indicaria sua exclusão, porém, devido a sua importância no contexto do mundo globalizado, optou-se por

considerar as sugestões dos juízes (J4, J7, J10) de mais bem definir e delimitar o termo internacionalização, assim como a indicação do seu emprego também aos Centros e IES (J6).

Nesse sentido, o numerador dos cálculos dos indicadores foi readequado com base nas 14 áreas de avaliação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes, que foram agrupadas em duas ações: mobilidade docente e discente (ativa e passiva) e cooperação internacional (acadêmica, científica e tecnológica)⁽²⁰⁾.

Tabela 2. Distribuição dos indicadores conforme o valor de CRV, na subdimensão Pesquisa, extensão, conhecimento e inovação, São Paulo-SP, 2018.

Indicadores		CVR
Subdimensão Pesquisa, extensão, conhecimento e inovação		
Fator Iniciação científica		
17	Taxa de docentes participantes em projetos de pesquisa	0.771
18	Taxa de estudantes participantes em projetos de pesquisa	0.657
19	Taxa de carga horária dedicada a projetos de pesquisa	0.200
Fator Programas de <i>stricto sensu</i>		
20	Conceito dos programas de mestrado/doutorado	0.886
21	Índice de publicação docente	0.457
22	Taxa de dissertações defendidas e aprovadas no programa de mestrado	0.571
23	Taxa de teses defendidas e aprovadas no programa de doutorado	0.543
Fator Programas, projetos e atividades de extensão		
24	Número de atividades de extensão voltadas à inserção comunitária	0.343
25	Taxa de estudantes com bolsa de extensão	0.457
26	Taxa de docentes envolvidos em atividades de extensão	0.457
27	Taxa de estudantes envolvidos em atividades de extensão	0.771
Fator Conhecimento e inovação		
28	Número de empresas incubadas	0.771
29	Taxa de projetos de inovação aprovados	0.800
30	Taxa de projetos de inovação executados	0.829
31	Taxa de patentes registradas	0.543

Quanto aos indicadores da Tabela 2, na subdimensão Programas *stricto sensu*, o indicador 21 obteve CVR de 0,457. A partir dos comentários de J6 e J10, procedeu-se à reformulação do indicador visando detectar a concentração de publicações por docente e segmentá-las por estrato *Qualis* CAPES, de forma que o cálculo ajustou-se na seguinte equação:

$$\frac{\sum \text{de artigos publicados por estrato} - \text{docente 1}}{\text{N}^\circ \text{ de docentes do programa}} = \frac{\sum \text{de artigos publicados por estrato} - \text{docente 2}}{\text{N}^\circ \text{ de docentes do programa}} = \dots = \frac{\sum \text{de artigos publicados por estrato} - \text{docente n}}{\text{N}^\circ \text{ de docentes do programa}}$$

Os indicadores 22 e 23 obtiveram um CRV de 0,571 e 0,543, respectivamente. O juiz J6 sugeriu determinar o índice de conformidade em 100%, o que foi acatado.

No Fator Programas, projetos e atividades de extensão, o indicador 24 obteve um CVR de 0,343. Os comentários (J5, J6, J7) referiam-se à imprecisão do descritor, que foi então reformulado para delimitar quais as atividades de extensão seriam consideradas, quais sejam: cursos, eventos, produção e publicação, e prestação de serviços, bem como as oito áreas temáticas da atividade extensionista: saúde, educação, trabalho, meio ambiente, comunicação, direitos humanos e justiça, tecnologia de produção e cultura⁽²¹⁾.

Para o indicador 25, o CVR obtido foi de 0,457, porém foi excluído em razão da alteração sugerida por J6, que está contemplada no indicador 27.

Para o indicador 26, obteve-se um CVR de 0,457. Os comentários dos juízes (J6, J7, J10) sugeriram melhorar a definição do termo extensão, o que foi aplicado também para os indicadores 24 e 26 como: práticas acadêmicas de caráter educativo,

cultural e científico, voltadas às necessidades da comunidade, que articulam o ensino e a pesquisa por meio de programas, projetos, cursos, eventos e prestação de serviço⁽²²⁾.

Para o Fator Conhecimento e inovação, o indicador 31 obteve um CRV de 0,543. O

especialista J6 recomendou definir melhor os itens a serem considerados no cálculo e, sendo assim, acrescentou-se à avaliação do indicador marcas, cultivares, desenho industrial, indicação geográfica, segredo industrial, conhecimentos tradicionais e topografia de circuito integrado⁽²³⁾.

Tabela 3. Distribuição dos indicadores conforme o valor de CRV, na subdimensão Gestão, São Paulo-SP, 2018.

Indicadores		CVR
Subdimensão Gestão		
Fator Gestão acadêmica		
32	Taxa de titulação docente	0.771
33	Taxa de carga horária docente em relação ao tipo de vínculo institucional	0.229
Fator Orçamento		
34	Despesas de custeio	0.286
35	Despesas com investimentos	0.286
36	Custo médio do curso de graduação	0.286
37	Custo médio do estudante de graduação	0.343
Fator Desenvolvimento humano		
38	Taxa de vagas não preenchidas de docentes/agentes universitários	0.486
39	Taxa de absenteísmo	0.086
40	Taxa de capacitação	0.543
Fator Infraestrutura		
41	Metros quadrados de salas de aula por estudante	0.686
42	Metros quadrados de auditórios por estudante	0.543
43	Número de sanitários disponíveis	0.686
44	Número de acervo bibliográfico básico	0.829
45	Número de acervo bibliográfico complementar	0.686
46	Assinaturas de periódicos especializados	0.457

Nota-se que, na Subdimensão Gestão, dos quatro indicadores que compõem o Fator Orçamento, três foram excluídos por apresentarem um CRV de 0,286 (34, 35, 36), restando Indicador 37, “Custo médio do estudante de graduação”, o qual apresentou CRV de 0,343.

A partir dos comentários dos especialistas (J2, J3, J4, J6) evidenciou-se a necessidade de melhorar a descrição do indicador, com definição das unidades de custo e detalhamento da fórmula. No entanto, apenas o indicador 37 não seria representativo do Fator Orçamento, sendo assim, decidiu-se pela exclusão deste.

Para o Fator Desenvolvimento Humano, o indicador 38 recebeu CVR de 0,486. Embora não tenha havido sugestões de como aprimorá-lo, optou-se por reformular o título do indicador para “Taxa de vagas do quadro de docentes”, de forma a delimitar melhor o objeto de interesse.

O indicador 40 obteve CVR de 0,543 e também não houve sugestões por parte dos juízes. Todavia, foi mantido, pois a preocupação com a qualificação do docente deriva também da sua relação direta com a qualidade da educação.

Quanto ao Fator Infraestrutura, o indicador 42, com CVR de 0,543, foi excluído frente ao argumento de J6 acerca da sua irrelevância, assim como o indicador 46 (CVR 0,457), para o qual também não houve sugestões dos juízes para aprimorá-lo.

DISCUSSÃO

A partir dos dados apresentados, constata-se que as Subdimensões e Fatores foram diferentemente valorados entre os juízes. Os indicadores que obtiveram maior número de contribuições dos especialistas foram os relativos a temáticas complexas, como a avaliação da evasão (5,9,13), internacionalização (15,16), publicação docente (21) e extensão (26).

Ao compararmos os achados evidenciados nas Tabelas de 1 a 3, nota-se que os menores valores de CRV foram obtidos na Subdimensão Gestão, principalmente no Fator Orçamento (CVR 0,300), o que pressupõe a existência de relação com o fato de este campo específico de saberes não ser, até então, contemplado nas dimensões avaliativas instituídas pelo Estado e, assim sendo, não percebido como

pertinente por parte dos avaliadores; no entanto, carece o desenvolvimento de outros estudos, tendo em vista elucidarem este fato.

Todas as recomendações dos especialistas foram acatadas e os indicadores com CVR inferiores a 0,300 foram excluídos, além dos indicadores 37, 42 e 46. Como não ocorreram discordâncias entre os juízes, a matriz de indicadores não necessitou nova etapa de validação.

Ao final dos procedimentos de validação de conteúdo, 38 indicadores obtiveram evidências de validade, dos quais 11 avaliam aspectos de estrutura, 10, de processos e 17, de resultados. Destes, 15 se referem à Subdimensão Ensino, 14, à Pesquisa, extensão, conhecimento e inovação, e oito, à Gestão⁽¹⁴⁾.

Na educação superior, assim como em toda organização de prestação de serviços, é imprescindível buscar maior eficiência nos processos, assim como maior eficácia em suas ações. A avaliação é uma ferramenta que instrumentaliza o gestor e que visa à excelência da qualidade do ensino^(6,9).

Avaliar não deve se limitar a mensurar resultados tendo em vista o controle, mas sim, formativamente subsidiar oportunidades de olhar para si de forma reflexiva, estabelecendo as inter-relações entre estruturas, processos e resultados, o que auxilia o processo decisório⁽⁶⁾.

No decorrer do desenvolvimento desta pesquisa, deparou-se com a dificuldade comumente encontrada em estudos de validação de instrumentos de medida, que perpassa a trajetória metodológica propriamente dita e culmina com os procedimentos de avaliação pelos juízes, sendo o número de especialistas uma limitação deste estudo, embora a cifra de 10 juízes atenda ao recomendado na literatura⁽²⁴⁾.

Outra limitação refere-se aos indicadores selecionados. Não houve a pretensão de esgotar todas as variáveis da dimensão Desempenho do construto qualidade da educação superior. A intenção foi propor um número de indicadores pertinentes para avaliar aspectos da estrutura, processos e resultados, válidos quanto ao seu conteúdo, de fácil aplicação e compreensão dos resultados, tendo em vista estarem formulados como

equações matemáticas com resultados expressos na forma de índices e taxas para o monitoramento e controle dos processos de autoavaliação institucional.

Quanto ao tipo de escala utilizada neste estudo, considera-se um avanço implementar a escala dicotômica para validação de métricas de qualidade. Comumente observa-se o emprego de escalas tipo *Likert* de quatro pontos para estudos desta natureza⁽¹²⁾; entretanto, a escala dicotômica é uma alternativa para garantir maior assimetria entre positividade e negatividade das respostas⁽²⁴⁾, sem repercussão na confiabilidade⁽²⁵⁾.

CONSIDERAÇÕES FINAIS

Na educação superior pública, assim como em todas as organizações, a avaliação proativa é imprescindível, pois permite o autoconhecimento a partir do estabelecimento da relação entre processos e resultados, possibilitando a detecção de potencialidades e fragilidades, base para o planejamento e a gestão.

O desenvolvimento desta pesquisa possibilitou obter evidências de validade de uma matriz de 38 indicadores quantitativos, formulados como equações matemáticas, que podem ser utilizados sistematicamente na avaliação interna institucional, complementando a avaliação qualitativa e subsidiando a identificação dos potenciais de melhoria, o que permite ser aplicado no âmbito dos Cursos, dos Centros de Estudo, Órgão e/ou Núcleos, respeitando-se as especificidades.

Não se trata de propor indicadores inéditos para a avaliação de desempenho de instituições públicas de ensino superior, mas sim, a partir dos processos avaliativos instituídos no país, propor uma matriz de indicadores considerados pertinentes por um painel de experts, para a avaliação interna e sistemática dessas organizações, tendo em vista o autoconhecimento.

Dando continuidade aos procedimentos metodológicos de validação de medidas, recomenda-se o desenvolvimento de novos estudos, visando testar a aplicabilidade e confiabilidade dos indicadores propostos.

METRICS VALIDATION TO EVALUATE QUALITY IN PUBLIC HIGHER EDUCATION INSTITUTIONS

ABSTRACT

Introduction: In public higher education, the complexity of evaluating goals to be reached and the uncertainties about metrics adopted by the country hinder the implementation of evaluation processes. **Objective:** To build and validate the content of quality indicators for evaluation of Higher Education Institutions (HEIs). **Method:** Methodological study developed in six stages, in which 10 HEI judges participated, between 2015 and 2018. The Delphi method was employed, and the Content Validity Ratio was calculated, with a value higher or equal to 0.62. A total of 46 indicators were proposed in the following subdimensions: Teaching (16), Research, Extension, Knowledge and Innovation (15), and Management (15). **Results:** Content validity evidence was found for 38 indicators, 11 of which related to structure, 10 to processes, and 17 to results, with the Management subdimension being the one with the lowest CVR values. **Further Considerations:** Validity evidence was found for a matrix of 38 quantitative indicators formulated as mathematical equations, which can be systematically applied to internal institutional evaluations to complement qualitative ones.

Keywords: Organization and Administration. Evaluation Studies as Topic. Indicators. Validation Studies. Higher Education.

VALIDACIÓN DE MÉTRICAS PARA EVALUAR LA CALIDAD EN INSTITUCIONES PÚBLICAS DE ENSEÑANZA SUPERIOR

RESUMEN

Introducción: En el ámbito de la educación superior pública, la complejidad para evaluar el alcance de los objetivos institucionales y las incertidumbres acerca de las métricas adoptadas por el Estado perjudican la implementación de procesos de evaluaciones. **Objetivo:** Construir y validar el contenido de indicadores de calidad para la evaluación de Instituciones Públicas de Enseñanza Superior (IES). **Método:** Estudio metodológico desarrollado en seis etapas, con la participación de 10 jueces de IES, entre 2015 y 2018. Se adoptó la Técnica Delphi y el cálculo del *Content Validity Ratio*, con valor mayor o igual a 0,62. Fueron propuestos 46 indicadores en las subdimensiones: Enseñanza (16); Investigación, extensión, conocimiento e innovación (15); y Gestión (15). **Resultados:** Se Constata la evidencia sobre la validez de contenido de 38 indicadores, de los cuales 11 de estructura, 10 de procesos y 17 de resultados, siendo la subdimensión Gestión la que obtuvo menores valores de CVR. **Consideraciones Finales:** Fueron obtenidas evidencias sobre la validez de una matriz de 38 indicadores cuantitativos, formulados como ecuaciones matemáticas, que pueden ser utilizados sistemáticamente en la evaluación interna institucional, complementando la evaluación cualitativa.

Palabras clave: Organización y administración. Estudios de Evaluación como Asunto. Indicadores. Estudios de validación. Educación superior.

REFERÊNCIAS

- Chaves SE, Cecim RB. Higher Education external evaluation in the health area: concerns and the size of margins. Interface (Botucatu). [on-line]. 2015 [citado em 28 jul. 2018]; 19(55):1233-1242. doi: <http://dx.doi.org/10.1590/1807-57622014.0728>.
- Andriola WB; Oliveira KRB. Auto avaliação institucional na Universidade Federal do Ceará (UFC): meio século de história. Avaliação. [on-line]. 2015 [citado em 02 fev 2019]; 20(2):489-512. Disponível em: <http://periodicos.uniso.br/ojs/index.php/avaliacao/article/view/1286>
- Figueiredo AC. Limites para afiliação à vida acadêmica de estudantes de camadas populares no contexto de expansão universitária. Educ. Pesqui. [on-line]. 2018. [citado em 02 fev 2019]; 44:1-18. doi: <http://dx.doi.org/10.1590/s1678-4634201844173462>.
- Morosini MC, Fernandes CB, Leite D, Franco MEDP, Cunha MI, Isaia SMA. Quality of higher education and the complex exercise of proposing indicators. Revista Brasileira de Educação [on-line]. 2016 jan-mar. [citado em 02 fev 2019]; 21(64):13-37. doi: <http://dx.doi.org/10.1590/S1413-24782016216402>.
- Ramos MY. Internationalization of graduate education in Brazil: rationale and mechanisms. Educ. Pesqui. [on-line]. 2018. [citado em 02 fev 2019]; 44:1-22. doi: <http://dx.doi.org/10.1590/s1517-9702201706161579>.
- Salsi MA, Silva DMG, Meirelles BHS. Evaluation in the Brazilian health system. Cienc Cuid Saude. [on-line]. 2018 [citado em 27 out 2018]; 17(2):1-6. doi: <http://dx.doi.org/10.4025/ciencucidsaude.v17i2.41937>.
- Barreyro GB, Rothen, JC. Percurso da avaliação da educação superior nos governos Lula. Educ. Pesqui. [on-line]. 2014 mar. [citado em 02 fev 2019]; 40(1):61-76. doi: <http://dx.doi.org/10.1590/S1517-97022014000100005>.
- Chaves SE, Cecim, RB. Higher education external evaluation in the health area: concerns and the size of margins. Interface (Botucatu) [on-line]. 2015. [citado em 02 fev 2019]; 19(55):1233-42. doi: <http://dx.doi.org/10.1590/1807-57622014.0728>.
- Trzesniak P. Indicadores quantitativos: como obter, avaliar, criticar e aperfeiçoar. Navus - Revista de Gestão e Tecnologia [on-line]. 2014 jul/dez. [citado em 02 fev 2019]; 4(2):05-18. doi: <https://doi.org/10.22279/navus.2014.v4n2.p5-18.223>.
- Afonso AJ. Protagonismos instáveis dos princípios de regulação e interfaces público/privado em educação. Educ. Soc. [on-line]. 2010 out/dez. [citado em 02 fev 2019]; 31(113):1137-56. Disponível em: <https://www.redalyc.org/articulo.oa?id=87315816005>.
- Polit DF, Beck CT. Essentials of nursing research: appraising evidence for nursing practice. 8th Edition, Lippincott Williams & Wilkins: Philadelphia; 2014.
- Coluci MZO, Alexandre NMC, Milani D. Construção de instrumentos de medida na área da saúde. Ciência & Saúde Coletiva. [on-line]. 2015. [citado em 12 ago 2018]; 20(3):925-36. doi: <http://dx.doi.org/10.1590/1413-81232015203.04332013>.
- Matitz QRS, Bulgacov S. O conceito desempenho em estudos organizacionais e estratégia: um modelo de análise multidimensional. RAC (Curitiba) [on-line]. 2011. [citado em 14 jul 2018]; 15(4): 580-607. doi: <http://dx.doi.org/10.1590/S1415-65552011000400003>.
- Zanin A, Poli OL, Moura GD, Jung CF, Caten CS. Definição de painel de indicadores de desempenho para instituições comunitárias de ensino superior. Revista GUAL [on-line] 2015 mai. [citado em 02 fev 2019]; 8(2):01-27. doi: <https://doi.org/10.5007/1983-4535.2015v8n2p1>.
- Uchoa CE. Elaboração de indicadores de desempenho institucional. Coordenação Geral de Programas de Capacitação /DDG. Brasília: ENAP/DDG, 2013. Disponível em: <http://repositorio.enap.gov.br/handle/1/2403>.
- Giannarou L, Zervas E. Using Delphi technique to build

consensus in practice. *Int. Journal of Business Science and Applied Management* [on-line]. 2014 [citado em 9 mai 2018]; 9(2): 66-82. Available from: <http://www.business-and-management.org/paper.php?id=106>.

17. Mophajan, H. Two criteria for good measurements in research: validity and reliability. *Annals of SpiruHaret University*. [on-line]. 2017 [citado em 15 jul 2018]; 17(3): 58-82. Available from: <https://mpira.ub.uni-muenchen.de/83458/>.

18. Lawshe, CH. Quantitative approach to content validity. *Personnel Psychology*. [on-line]. 1975 [citado em 14 mai 2018]; 28:563-75. doi: <http://dx.doi.org/10.1111/j.1744-6570.1975.tb01393.x>.

19. Wilson FR, Pan W, Schumsky DA. Recalculation of the critical values for Lawshe's Content Validity Ratio. *Measurement and evaluation in counseling and development*. [on-line]. 2012 [citado em 12 mai 2018]; 45(3):197-210. doi: <https://doi.org/10.1177/0013164412440286>.

20. Ministério da Educação. CAPES. A internacionalização na universidade brasileira: resultados do questionário aplicado pela Capes. Diretoria de Relações Internacionais. [on-line]. 2017 [citado em 8 ago 2018]. Disponível em: <https://www.capes.gov.br/images/stories/download/diversos/A-internacionalizacao-nas-IES-brasileiras.pdf>.

21. Paula JA. A extensão universitária: história, conceito e

propostas. *Interfaces - Revista de Extensão*. [on-line]. 2013 [citado em 9 ago 2018]; 1(1):05-23. Disponível em: <https://www.ufmg.br/proex/revistainterfaces/index.php/IREXT/article/view/5>.

22. Nogueira MDP. O fórum de pró-reitores de extensão das universidades públicas brasileiras: um ator social em construção. *Interfaces - Revista de Extensão* [on-line]. 2013 [citado em 14 ago 2018]; 1(1): 35-47. Disponível em: <https://www.ufmg.br/proex/revistainterfaces/index.php/IREXT/article/view/7>.

23. Silva EBF, Coelho MRS, Santos MVV, Leite RS. Pesquisa e inovação: a propriedade intelectual do estado de Minas Gerais. *Cadernos de Prospecção* [on-line]. 2018 set. [citado em 02 fev 2019]; 11(3):757-69. doi: <http://dx.doi.org/10.9771/cp.v11i3.26966>.

24. Haynes SN, Richerd DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. *Psychological Assessment*. [on-line]. 1995 [citado em 9 ago 2018]; 7(3):238-47. doi: <http://dx.doi.org/10.1037/1040-3590.7.3.238>.

25. Lucian R, Dornelas JS. Mensuração de atitude: proposição de um protocolo de elaboração de escalas. *RAC (Rio de Janeiro)*. [on-line]. 2015 [citado em 8 ago 2018]; 19(3):157-77. doi: <http://dx.doi.org/10.1590/1982-7849rac20151559>.

Endereço para correspondência: Dagmar Willamowius Vituri. Rua Virgílio Jorge nº 695. Jardim San Remo. CEP 86062-270, Londrina-PR. E-mail: dagmar@uel.br

Data de recebimento: 26/10/2018

Data de aprovação: 06/02/2019