

# Avaliação do Docente pelo Discente no Âmbito do Ensino Técnico Integrado: evidências de validade da Escala SIR-II

▮ Roberto Brasileiro Paixao \*

▮ Anamaria Azevedo Lafeta Rabelo \*\*

▮ Adriano Leal Bruni \*\*\*

---

## Resumo

Este artigo teve como objetivo fazer a tradução, retrotradução, revisão por especialistas e pré-testagem da segunda versão da escala *Student Instructional Report* (SIR-II) da organização *Educational Testing Service* (ETS), especificamente na parte em que trata da avaliação do docente. O instrumento analisado é composto por cinco dimensões: Organização e Planejamento; Comunicação; Interação Professor/Aluno; Avaliações, Exames e Notas; e Métodos Suplementares de Ensino. A amostra do presente estudo foi composta por 5.622 casos gerados por estudantes de cursos técnicos integrados ao ensino médio de diferentes *campi* do Instituto Federal de Educação, Ciência e Tecnologia do Norte de Minas Gerais (IFNMG). O modelo demonstrou ajustamento quanto à unidimensionalidade, confiabilidade e convergência, apresentando evidências de validade da escala em sua versão brasileira.

**Palavras-chave:** Avaliação. Escala. Tradução. Validação.

---

\* Doutor em Administração pela Universidade Federal da Bahia. Professor do Núcleo de Pós-graduação em Administração da Universidade Federal da Bahia; E-mail: robertobrazileiro@gmail.com.

\*\* Mestre em Administração pela Universidade Federal da Bahia. Assessora de Pesquisa Institucional do Instituto Federal de Educação, Ciência e Tecnologia do Norte de Minas Gerais. E-mail: anatotty@hotmail.com.

\*\*\* Doutor em Administração, Universidade de São Paulo. Professor do Núcleo de Pós-graduação em Administração da Universidade Federal da Bahia. E-mail: albruni@gmail.com.

## 1 Introdução

Conforme expresso em sua lei de criação, Lei nº 11.892 de 29 de dezembro de 2008, os Institutos Federais de Educação, Ciência e Tecnologia devem fornecer educação buscando a formação profissional e a qualificação dos seus estudantes para atuar no mundo do trabalho (BRASIL, 2008). Para garantir o alcance desse e de outros objetivos presentes na Lei citada, os Institutos Federais devem avaliar a instituição e seus docentes, de forma a ofertar uma educação de qualidade. Sousa (2008) acredita que os gestores educacionais apostam na avaliação como instrumento de melhoria da qualidade de ensino.

Ao se analisar a qualidade do ensino ofertado, um dos indicadores avaliados é o desempenho do corpo docente. Esse tipo de avaliação é realizado principalmente pelas Comissões Próprias de Avaliação (CPA), como exigência da Lei nº 10.861, de 14 de abril de 2004, a qual discorre que as avaliações das instituições de ensino superior deverão avaliar obrigatoriamente o desenvolvimento profissional do docente (BRASIL, 2004). Dentre os objetivos da CPA tem-se a identificação das condições de ensino oferecidas aos estudantes, em especial as relativas ao perfil do corpo docente, às instalações físicas e à organização didático-pedagógica.

Essa avaliação docente pode ser efetivada pelos vários envolvidos no processo de oferta da educação: comunidade, egressos, docentes, técnicos administrativos em educação, discentes, dentre outros. Nesse trabalho, aborda-se a visão dos discentes sobre a avaliação dos docentes.

Na literatura nacional foram encontradas várias escalas que tratam da avaliação docente pelo discente (AVRICHIR; DEWES, 2006; MUSETTI, 2007; GOMES; BORGES, 2008; BITTENCOURT et. al., 2011; MACEDO, 2011; SOUZA; CHISSOLUCOMBE, 2012; FREITAS et al., 2014; SILVA et al., 2017). Algumas delas tratam de dimensões similares, apesar de terem públicos distintos.

Buscou-se, neste estudo, uma escala que pudesse ser utilizada para alunos de cursos técnicos, com ensino médio integrado ou subsequente, e que não fosse tão longa a ponto de inviabilizar sua aplicação em sala de aula, que abrangesse múltiplas dimensões do processo de avaliação docente e que tivesse ampla utilização internacional. O instrumento *Student Instructional Report* (SIR versão II ou apenas SIR-II) da organização

*Educational Testing Service* (CENTRA, 2007) apresenta essas características e, por isso, foi escolhido para ser utilizado.

Entretanto, para a utilização desse instrumento, faz-se necessário realizar o processo de adaptação cultural da escala, que envolve: a) a tradução; b) a retrotradução; c) a revisão dos itens; d) o pré-teste; e e) a possível ponderação de valores em relação ao contexto cultural (GUILLEMIN; BOMBARDIER; BEATON, 1993; HILL; HILL, 2002). O objetivo do presente estudo foi realizar as quatro primeiras etapas, ou seja, fazer a tradução-retrotradução, revisão e pré-testagem da escala SIR-II, especificamente na parte em que trata da avaliação docente.

## **2 Avaliação do docente pelo discente**

Na maioria das instituições de ensino os questionários de avaliação são aplicados pela CPA (Comissão Própria de Avaliação). Montemór (2013) argumenta que esse tipo de avaliação é realizada com a participação de pessoas da instituição, mas que não são necessariamente os próprios professores.

Os instrumentos de avaliação docente em sua maioria, abordam os temas propostos pela Lei nº 10.861, de 14 de abril de 2004, por meio de diferentes aspectos, como: planejamento e gestão das atividades de ensino; didática adotada pelo professor no desenvolvimento das atividades de ensino; uso adequado dos materiais e procedimentos didáticos; avaliação adotada pelo professor para verificar o aprendizado discente; comunicação e interação com os alunos, definida pelo contato e envolvimento do aluno e do professor; clareza e objetividade na apresentação e desenvolvimento do conteúdo ministrado; estímulo ao questionamento; pontualidade e assiduidade (BRAGA; MONTEIRO, 2005; ANDRIOLA, 2011; BOUTH, 2013).

Interessante notar que as questões (ou itens) sobre avaliação docente pelo discente que apresentam maior convergência entre os autores são: repassa informação adequadamente/clareza, abre espaço para sugestão de aluno e incentiva participação crítica, relaciona teoria e prática e reflete sobre os resultados da disciplina. Todos eles estão dispostos na escala SIR-II, escolhida para ser utilizada no presente estudo. O quadro 1 apresenta os itens abordadas pelos autores pesquisados, seja quando a escala foi disponibilizada na íntegra ou quando os itens foram captados a partir das análises dos respectivos resultados.

Quadro 1 - Itens constantes da avaliação do docente pelo discente encontradas nos trabalhos pesquisados

Itens	Trabalhos Pesquisados								
	Macedo (2001)	Avrichir ; Dewes (2006)	ETS (2017)	Mu-setti (2007)	Gomes e Borges (2008)	Bitten-court et al. (2011)	Souza; Chissoluc ombe (2012)	Freitas et al. (2014)	Silva et al. (2017)
Repassa Informação adequadamente/clareza	X	X	X	X	X	X	X		X
Uso de recursos didáticos/adequação		X	X	X		X	X	X	X
Aplica exercícios de fixação/revisão e qualidade do material			X				X		
Abre espaço para sugestão de aluno e incentiva participação	X	X	X	X	X		X	X	X
A linguagem e a atualização das fontes utilizadas	X		X	X	X	X	X		X
Relaciona teoria e prática/uso de exemplos	X		X	X	X		X		X
Incentiva atividades extraclases					X	X	X	X	
Pontualidade/Assiduidade	X		X				X	X	X
Apresenta/ segue o plano de ensino	X		X	X	X	X	X		
Reflexão sobre os resultados da disciplina	X	X	X	X		X		X	X
Aproveitamento do tempo da aula			X	X	X	X			
Motivação para aprender mais			X		X	X		X	
Contribuição para um ambiente favorável de aprendizagem			X	X	X	X			
Clareza dos critérios de avaliação	X	X	X		X			X	
Mantém a atenção do aluno na aula		X	X						X
Postura crítica e entusiasmo quanto aos conteúdos que leciona		X	X						X
O respeito do professor pelos alunos			X						

Continua

Itens	Trabalhos Pesquisados									Conclusão
	Macedo (2001)	Avrichir ; Dewes (2006)	ETS (2017)	Mu-setti (2007)	Gomes e Borges (2008)	Bitten-court et al. (2011)	Souza; Chissoluc ombe (2012)	Freitas et al. (2014)	Silva et al. (2017)	
A disponibilidade em atender o aluno			X						X	
Desenvolvimento de atividades em grupo			X							
As avaliações cobram os aspectos importantes do curso			X							

Fonte: Os autores (2017).

Moreira (1981) e Tejedor e Monteiro (1990) consideram a avaliação feita pelo discente como a principal fonte de informação sobre a atuação do docente. O aluno, como bem apropriada-se Brandalise (2012), deve sentir-se sujeito da avaliação, ele não pode sentir a avaliação como algo externo.

Para Bouth (2013, p. 377), a avaliação docente feita pelo discente é necessária para garantir qualidade do processo educacional. Os autores recomendam a utilização de alguma escala, na qual conste elementos que indiquem minimamente a qualidade na educação.

## 2.1 Tradução e Adaptação Cultural de Escalas

Hill e Hill (2002) apontam soluções para a utilização de escalas ainda não traduzidas para o idioma dos respondentes. O processo inicia-se com a tradução-retrotradução, ou seja, a escala é traduzida e, na sequência, retraduzida para o idioma original. Posteriormente tem-se a validação analítica dos itens até alcançar a etapa de pré-testagem.

Conforme Beaton et al. (2000) e Fetter et al. (2014) o termo adaptação cultural refere-se a um processo para alcançar a equivalência semântica entre a escala original e a adaptada, de forma que o instrumento final produza efeito semelhante ao original.

Hill e Hill (2002) entendem que o processo de tradução literal de uma escala não garante a compreensão dos termos, sendo importante realizar as adaptações de semântica, percepções e de culturas entre os diferentes idiomas.

De acordo com Jesus e Rowe (2014), um dos objetivos da validação linguística é que o instrumento seja facilmente compreendido entre os respondentes.

Esse processo, iniciado com a tradução dos itens da escala para o idioma em que se pretende aplicá-lo, é essencial para a aplicação do instrumento, ainda que estes já tenham sido testados na versão original (FETTER et al., 2014).

Este processo deve obedecer algumas etapas para a validação da escala. A primeira etapa é a tradução e validação da tradução por pessoas com conhecimentos sólidos e domínio da língua da versão original do instrumento, além de experiência na tradução de escalas (GUILLEMIN; BOMBARDIER; BEATON, 1993; HILL; HILL, 2002). Beaton et al. (2000) destaca a importância de a tradução ser feita por pessoas com fluência em ambos os idiomas, mas principalmente ser nativo no idioma-alvo.

A segunda etapa constitui-se de verificação da tradução através da retrotradução e comparação com o original (GUILLEMIN; BOMBARDIER; BEATON, 1993). Esta etapa visa garantir que haja equivalência semântica, idiomática, experiencial e conceitual (BORSA; DAMÁSIO; BANDEIRA, 2012).

A terceira etapa revela-se pela validação conceitual por meio da análise realizada por especialistas na área, denominado comitê de revisão (GUILLEMIN; BOMBARDIER; BEATON, 1993). Além das questões ligadas a estrutura, *layout*, instruções do instrumento e abrangência, os especialistas observam principalmente se as expressões podem ser generalizadas para diferentes culturas (BORSA; DAMÁSIO; BANDEIRA, 2012).

A quarta etapa refere-se ao estudo piloto ou pré-testagem para a verificação da clareza e compreensão dos itens, onde os testes de unidimensionalidade, confiabilidade e convergência são realizados (GUILLEMIN; BOMBARDIER; BEATON, 1993; HILL; HILL, 2002).

Uma quinta etapa refere-se à ponderação dos escores em relação ao contexto cultural em que o instrumento será aplicado (GUILLEMIN; BOMBARDIER; BEATON, 1993). Esta fase não está contemplada nesta pesquisa.

Como não se objetiva nesse estudo exaurir os diversos métodos existentes para a tradução de escalas, optou-se por utilizar o método apresentado por Hill e Hill (2002). Cabe ressaltar que existem outros métodos de tradução de escalas, tais como o método proposto por Herdman, Fox-Rushby e Badia (1998), o método de *MAPI Research Institute*, similar ao proposto por Beaton et al. (2000), o método de Swaine-Verdier et al. (2004),

citados por Jesus e Rowe (2014), o método apresentado por Fetter et al. (2014), a metodologia proposta pela Organização Mundial de Saúde apresentada por Oliveira, Hildenbrand e Lucena (2015) ou a abordagem proposta por Borsa, Damásio e Bandeira (2012).

### 3 Procedimentos metodológicos

A avaliação docente proposta nesse trabalho relaciona-se às atitudes dos docentes em seu cotidiano escolar e foi feita a partir da escala de avaliação docente da escala SIR-II. Os alunos avaliaram os docentes nos quesitos referentes ao planejamento e organização da disciplina, comunicação do professor, interação entre professor e aluno, avaliações dadas pelos professores e métodos e materiais didáticos utilizados pelos docentes. No decorrer desse tópico serão apresentados os procedimentos metodológicos utilizados para a validação da escala.

#### 3.1 Descrição da amostra

O instrumento de coleta foi aplicado de maneira presencial na sala de aula do aluno, de modo a alcançar o maior número de respondentes possíveis. A coleta ocorreu entre os meses de novembro de 2016 a fevereiro de 2017, nos *Campi* do IFNMG das cidades de Salinas, Araçuaí, Almenara, Januária, Arinos, Montes Claros e Pirapora. Houve aprovação das instâncias superiores (Diretores de unidade e Pró-reitor de Pesquisa) do IFNMG para a aplicação da pesquisa, uma vez que o instituto não tinha à época um Comitê de Ética constituído.

A amostra pode ser caracterizada como não probabilística e por conveniência. O instrumento de coleta foi distribuído para 2.104 alunos, conforme Tabela 1.

Tabela 1 - Alunos Pesquisados do IFNMG no período de 2016 – 2017

CAMPUS	TURMAS	Nº DE ALUNOS	Nº DE RESPONDENTES
Almenara	1º, 2º e 3º Séries dos cursos Técnico em Agropecuária, Técnico em Informática e Técnico em Zootecnia, Integrados ao Ensino Médio	287	63
Araçuaí	1º, 2º e 3º Séries dos cursos Técnico em Informática e Técnico em Meio Ambiente, integrados ao Ensino Médio. 1º Série de Técnico em Agroecologia e Técnico em Agrimensura Integrados ao Ensino Médio.	277	128

Continua

			Conclusão
CAMPUS	TURMAS	Nº DE ALUNOS	Nº DE RESPONDENTES
Arinos	1º, 2º e 3º Séries dos cursos Técnico em Agropecuária, Técnico em Informática e Técnico em Meio Ambiente, Integrados ao Ensino Médio	381	172
Januária	1º, 2º e 3º Séries de Técnico em Informática, Técnico em Agropecuária e Técnico em Meio Ambiente, Integrados ao Ensino Médio	396	139
Montes Claros	1º, 2º e 3º Séries de Técnico em Informática e Técnico em Química, Integrados ao Ensino Médio	174	61
Pirapora	1º, 2º e 3º Séries de Técnico em Informática e Técnico em Edificações, Integrados ao Ensino Médio	171	70
Salinas	1º, 2º e 3º Séries de Técnico em Agropecuária, Técnico em Agroindústria e Técnico em Informática, Integrados ao Ensino Médio	418	170
Total		2.104	803

Fonte: Os autores (2017).

### 3.2 Descrição do instrumento

O instrumento SIR-II deriva do questionário original SIR, aplicado para alunos dos Estados Unidos nos últimos vinte e cinco anos (CENTRA; GAUBATZ, 2000). No Brasil ainda não há uma versão em português validada. Os itens constantes no instrumento estão disponíveis em alguns trabalhos científicos, como Centra e Gaubatz (2000) e Centra (2007).

O SIR-II é composto de oito escalas: Organização e planejamento (5 itens); Comunicação (5 itens); Interação professor/aluno (5 itens); Avaliações, exames e notas (6 itens); Métodos suplementares de ensino (7 itens); Resultados do curso (5 itens); Envolvimento e esforço do estudante (3 itens); e Dificuldade, carga de trabalho e ritmo do curso (3 itens). O detalhamento dos itens é apresentado no quadro 2.

Para avaliar professores dos cursos técnicos integrados ao ensino médio do Instituto Federal do Norte de Minas Gerais (IFNMG) foram utilizadas apenas as questões relacionadas à docência, limitando o instrumento às cinco primeiras dimensões apresentadas. Foram retiradas duas dimensões e, nas dimensões restantes, foram retirados três itens que se referem exclusivamente ao ensino universitário ou a disciplinas isoladas, não sendo aplicável aos estudantes e às disciplinas do ensino médio.



As dimensões Envolvimento e esforço do estudante e Dificuldade, carga de trabalho e ritmo do curso não foram usadas porque não tratam da avaliação do docente, mas sim da autoavaliação do discente. O item 23 – Trabalho de conclusão de curso ou projeto foi excluído porque não há no ensino médio a atividade de trabalho de conclusão de curso. O item 26 – Uso de estudos de casos, simulações ou dramatização foi excluído porque não há no ensino médio atividades com estudos de casos, simulações ou dramatização. E o item 27 – Diários de classe ou registros exigidos dos alunos também foi excluído por não haver tal exigência no IFNMG. Todos os itens excluídos fazem parte da dimensão Métodos suplementares de ensino, a qual neste trabalho teve quatro itens ao final. Ressalta-se que a Educational Testing Service (2017) informa sobre a possibilidade de flexibilidade no uso do instrumento, ou seja, de excluir alguma dimensão ou item do instrumento.

Optou-se por utilizar o SIR-II devido: a) à amplitude de dimensões analisadas; b) à menor extensão do instrumento, se comparado aos demais, sendo apropriado para aplicação presencial em sala de aula; e c) à validação internacional, sendo um dos instrumentos mais usados nos Estados Unidos e Canadá (CASHIN, 1990; CENTRA, 2007; SERRANO; VIZCARRA, 2013), o que também demonstra sua facilidade de entendimento por parte dos respondentes.

Segundo a Educational Testing Service (2017), o feedback abrangente fornecido nos relatórios SIR-II é útil para professores e gestores educacionais, que podem utilizar os dados obtidos para avaliar o curso; identificar pontos fortes de ensino e oportunidades de melhoria em disciplinas, cursos e aulas; informar decisões administrativas; e fornecer informações aos alunos sobre a qualidade da educação.

### **3.3 Processo de Tradução-retrotradução**

O processo de tradução da escala na versão em inglês para o português do Brasil deu-se utilizando procedimentos semelhantes ao método tradução-retrotradução apresentado por Hill e Hill (2002). Dessa forma buscou-se evidências de que a versão adaptada para a língua portuguesa é equivalente à sua versão original no idioma inglês.

O primeiro passo consistiu em traduzir literalmente o SIR-II do inglês para o português brasileiro, feito por um dos autores do trabalho. Em função da limitação orçamentária da pesquisa, o processo de tradução foi feito com apenas um profissional, quando a

recomendação é que sejam pelo menos dois para que haja comparação (HILL; HILL, 2002; BORSA; DAMÁSIO; BANDEIRA, 2012).

O segundo passo foi submeter os itens traduzidos a um professor do curso de Administração do IFNMG Campus Avançado Janaúba, que fez as suas considerações e contribuições sobre o entendimento e compreensão das questões. De forma que o item 3 traduzida inicialmente para “O comando do instrutor sobre o assunto”, passou a ser “Orientações do instrutor sobre os conteúdos”. O item 6 passou de “a capacidade do instrutor de tornar clara e compreensível as apresentações” para “a capacidade do professor de tornar clara e compreensível a aula”. O “comando do instrutor de inglês falado (ou a língua usada no curso)” do item 7 passou a ser “usa o idioma da aula de forma clara/ Usa linguagem clara”. O item 8 foi adaptado de “o uso de exemplos ou ilustrações pelo instrutor para esclarecer o curso material” para “o uso de exemplos ou ilustrações pelo professor para esclarecer o material do curso”. O item 11 passou de “A disponibilidade e a capacidade de resposta dos alunos” para “A habilidade do instrutor e a capacidade de responder aos alunos”. E o item 20 foi adaptado de “A qualidade global do(s) livro(s)” para a “Qualidade do material didático utilizado”.

Em seguida, a versão foi submetida a uma professora de Informática fluente em inglês do Campus Avançado Porteirinha e a um professor formado em Letras com habilitação em Letras e ao idioma Inglês para que fosse retrotraduzido, seguindo as adaptações do outro professor. Foram encontradas significativas diferenças e algumas adaptações necessárias a serem feitas de modo a melhorar a clareza, a pertinência e a equivalência dos itens traduzido. Decidiu-se por trocar a palavra instrutor em todas as questões para professor. A expressão “idioma da aula” foi substituída para “linguagem da aula”, já que a versão original prioriza a aula dada no idioma inglês. A expressão “fora do horário de aula” também foi ponto de confronto, já que os professores preferiam a palavra extraclasse para definir a expressão. O termo “exercícios de laboratório” tal como disposto no item 24, foi substituído por “exercícios práticos”.

Esses conflitos encontrados na retrotradução foram sanados em uma reunião entre um dos autores e os três professores. A versão original americana foi comparada com a retrotradução, para a verificação se algumas das mudanças sugeria alteração no conteúdo dos itens. Nessa reunião, após a geração de três versões para o texto final dos itens, alcançou-se uma versão de consenso.

Antes da aplicação do instrumento, um profissional Técnico em Assuntos Educacionais, especialista em educação verificou a sua adequação o que levou ao consenso final entre os tradutores, os autores e o especialista, culminando na versão de teste, a qual foi aplicada a cinco alunos do IFNMG – Campus Avançado Porteirinha, do curso técnico em Informática.

Uma das considerações feitas pelos alunos refere-se ao item 1, a expressão “requisitos do curso” gerou alguma dúvida por parte de três dos cinco alunos. Os cursos do IFNMG não apresentam pré-requisitos. Em consenso com os professores da área de Administração e da área de Informática, a expressão foi substituída por “objetivos da disciplina”.

O item 25 sobre uso de computador como auxiliar na instrução também suscitou dúvidas nos alunos. Devido a justificativa dos mesmos, estudantes do curso Técnico em Informática, o termo computador é muito restrito, o que pode sugerir dificuldades nas respostas dos alunos. Sendo que professores de outras áreas, como de Geografia, podem não utilizar computador, mas outras tecnologias em suas aulas. Assim, o item foi ampliado para “o uso de tecnologias para ajudar na compreensão da aula”.

Após esse procedimento foi possível verificar as equivalências: semântica, palavras com o mesmo significado; idiomática, uso de expressões equivalentes; experiencial, uso de itens que são aplicáveis na outra cultura; e conceitual, termos que avaliem o mesmo aspecto (BORSA; DAMÁSIO; BANDEIRA, 2012). Os itens originais e suas respectivas versões finais estão dispostos no quadro 2.

Quadro 2 - Versão original e versão final da escala SIR-II

Versão Original	Versão Final
Organização e planejamento	
1. “The instructor's explanation of course requirements”.	1. Explicação do professor sobre os objetivos da disciplina.
2. “The instructor's preparation for each class period”.	2 Preparação do professor para cada período de aula.
3. “The instructor's command of the subject matter”.	3 Explicações do professor sobre a matéria.
4. “The instructor's use of class time”.	4 O uso do tempo de aula pelo professor.
5. “The instructor's way of summarizing or emphasizing important points in class”.	5 A maneira do professor resumir ou enfatizar os pontos importantes na aula.

Continua

Continuação

Versão Original	Versão Final
<b>Comunicação</b>	
6. "The instructor's ability to make clear and understandable presentations".	6 A capacidade do professor de tornar clara e compreensível a aula
7. "The instructor's command of spoken English (or the language used in the course)".	7 A linguagem utilizada pelo professor para ministrar as aulas.
8. "The instructor's use of examples or illustrations to clarify course material".	8 Uso de exemplos ou ilustrações pelo professor para esclarecer a matéria
9. "The instructor's use of challenging questions or problems".	9 O professor propõe perguntas ou problemas desafiadores.
10. "The instructor's enthusiasm for the course material".	10 O entusiasmo do professor pelo material do curso.
<b>Interação professor/aluno</b>	
11. "The instructor's helpfulness and responsiveness to students".	11 A capacidade do professor em auxiliar e responder aos alunos.
12. "The instructor's respect for students".	12 O respeito do professor pelos alunos.
13. "The instructor's concern for student progress".	13 A preocupação do professor com o progresso do aluno.
14. "The availability of extra help for this class (taking into account the size of the class)".	14 A disponibilidade em atender o aluno extraclasse.
15. "The instructor's willingness to listen to student questions and opinions".	15 A disposição do professor em ouvir as perguntas e opiniões dos alunos.
<b>Avaliações, exames e notas</b>	
16. "The information given to students about how they would be graded".	16 As informações dadas aos alunos sobre os critérios de avaliação.
17. "The clarity of exam questions".	17 A clareza das perguntas da avaliação.
18. "The exams' coverage of important aspects of the course".	18 As avaliações cobram os aspectos importantes do curso.
19. "The instructor's comments on assignments and exams".	19 Os comentários do professor sobre as atividades e avaliações.
20. "The overall quality of the textbook(s)".	20 A qualidade do material didático utilizado.
21. "The helpfulness of assignments in understanding course material".	21 As atividades aplicadas ajudam na compreensão do curso.
<b>Métodos suplementares de ensino</b>	
22. "Problems or questions presented by the instructor for small group discussions".	22 Os problemas ou perguntas apresentadas pelo professor para discussões.
23. "Term paper(s) or project(s)".	Suprimida
24. "Laboratory exercises for understanding important course concepts".	23 Exercícios práticos para a compreensão dos importantes conceitos do curso.

Continua

		Conclusão
Versão Original	Versão Final	
Métodos suplementares de ensino		
25. "Assigned projects in which students worked together".	24. Desenvolvimento de atividades em grupo	
26. "Case studies, simulations, or role playing".	Suprimida	
27. "Course journals or logs required of students".	Suprimida	
28. "Instructor's use of computers as aids in instruction".	25. Uso de tecnologias para ajudar na aula.	

Fonte: Os autores (2017).

### 3.4 Procedimentos de Análise

Para a verificação de evidências de validade do SIR-II optou-se por utilizar os procedimentos de Hair Jr. *et al.* (2005) e Lima Filho e Bruni (2013), os quais citam as verificações sobre unidimensionalidade, confiabilidade e convergência.

A validade no entendimento de Hair Jr. *et al.* (2005, p. 112) é "o grau em que uma escala ou conjunto de medidas representa o conceito de interesse". Para a validade, são aceitas as regras de validação dispostas no quadro 3. Convém ressaltar que há recomendação para que o termo a ser utilizado seja evidencias de validade (AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, 2014).

Quadro 3 - Regras de validação de uma escala

Atributos	Técnicas ou Estatísticas	Regras para validação
1. Dimensionalidade	Análise Fatorial Exploratória	Apenas um autovalor deve existir na escala.
	Índice de KMO	Maior que 0,7: Desejável. Menor que 0,5: Inaceitável.
	Teste de esfericidade de Bartlett	Deve ser significativo.
2. Confiabilidade	Alfa de Cronbach	Alfa de Cronbach > 0,70.
3. Convergência	Coeficiente de Pearson	Coeficientes de Pearson > 0.

Fonte: LIMA FILHO; BRUNI (2013, p. 233).

Sobre a dimensionalidade, Hair Junior. *et al.* (2005) apresentam que os dados devem ser unidimensionais, de maneira a representar um único conceito entre os itens. A análise fatorial exploratória fornece ao pesquisador uma ferramenta útil para visualizar conjuntos de variáveis relacionadas e simplificar as outras estruturas de variáveis.

O índice de Kaiser Meyer Olkin (KMO) e o teste de esfericidade de Bartlett servem para testar a adequação da amostra levantada para utilização da análise fatorial (HAIR JR. et al., 2005; LIMA FILHO; BRUNI, 2013). Dessa forma, quando o KMO for superior a 0,7 e o teste de Bartlett for significativo, a realização da análise fatorial é apropriada.

A confiabilidade indica o grau de consistência entre as múltiplas medidas de uma variável, ou seja, o modo como é medido a variável. Os indicadores de construtos altamente confiáveis são também altamente inter-relacionados. Para a medida diagnóstica de confiabilidade de uma escala utiliza-se o Alfa de Cronbach, que deve ser um limite mínimo de 0,60. Sendo que as medidas mais próximas a 1 indicam maior confiabilidade entre os indicadores (HAIR JUNIOR. et al., 2005).

A convergência apresenta o grau em que duas variáveis são correlacionadas (HAIR JUNIOR. et al., 2005). Para tanto, Dancey e Reidy (2006) apresentam o teste de Coeficiente de Pearson ou Coeficiente da correlação momento-produto. Quando o coeficiente é igual a zero, tem-se que não existe nenhuma relação entre as variáveis. Se igual a 1, a relação entre as variáveis é perfeita. Espera-se que itens de uma mesma dimensão estejam com correlações mais fortes.

Em resumo, o pesquisador deve estabelecer conceitual e empiricamente que os seus constructos são válidos, unidimensionais e confiáveis (HAIR JUNIOR. et al., 2005).

#### **4 Resultados e Discussão**

Do total de 2.104 alunos, 803 responderam o instrumento de coleta, após terem concordado com o Termo de Assentimento, para os menores de idade ou de Termo de Consentimento, para os maiores de idade. Os alunos menores foram autorizados por seus pais e responsáveis por meio do Termo de Consentimento Livre e Esclarecido.

Dos 803 questionados gerou-se um banco de dados de 5.622 registros, já que os alunos responderam questões sobre sete disciplinas do núcleo comum, relacionadas a todas as séries e cursos do ensino médio integrado, oferecidas pelo IFNMG, quais sejam português, matemática, história, geografia, química, física e biologia. E também foram avaliados 125 professores efetivos e contratados da Instituição. Dessa forma, a análise sobre a validação das escalas foi realizada a partir desses 5.622 casos.

De acordo com a análise descritiva, a amostra apresentou ser em sua maioria do gênero feminino (54,5%), ter entre 16 e 18 anos (85,3%), não ser repetente (91,4%) e ingressou na Instituição através da ampla concorrência (60,77%).

A escala SIR-II foi validada em todas as dimensões, apresentando valores significativos quanto a dimensionalidade, confiabilidade e convergência, conforme os testes realizados com os 5.622 casos.

Os testes realizados para validar as cinco dimensões da avaliação docente quanto a dimensionalidade, confiabilidade e convergência (correlação) são apresentados na tabela 2.

Tabela 2 - Validação das Dimensões das Escalas de Avaliação docente

Atributos / Dimensões	Técnicas	Organização e Planejamento	Comunicação	Interação	Avaliação	Métodos
Dimensão- nalidade	An. Fatorial Exploratória	Único autovalor em todas as dimensões				
	Teste Bartlett	p-valor <0,001 em todas as dimensões				
	KMO	0,879	0,861	0,868	0,884	0,800
Confiabilidade	Alfa de Cronbach	0,902	0,863	0,859	0,864	0,822
Convergência	Coefficiente de Pearson	Todos positivos e significativos				

Fonte: Os autores (2017).

A medida de Kaiser-Meyer-Olkin verificou a adequação amostral para a análise de todas as dimensões. Todos os testes de KMO foram maiores que 0,7.

O teste de esfericidade de Bartlett indica que as correlações entre os itens são desejáveis para a realização da análise fatorial. Todas as dimensões apresentaram nível de significância aproximadamente nulo (p-valor < 0,001).

A partir da análise fatorial exploratória pôde-se verificar que o fator 1 explica 65,2% da variância total. Já o fator 2 explica 56,62% da variância total. Os fatores 3, 4 e 5 explicam respectivamente, 56,04%, 51,57% e 54,19% das variâncias totais.

Os coeficientes Cronbach das dimensões indicaram adequada confiabilidade. Todos os valores obtidos foram maiores que 0,70. A dimensão de maior confiabilidade é a primeira (0,902). Por outro lado, a menor é a dimensão 5, com valor de 0,822. O alfa de

Cronbach de cada uma das dimensões foi considerado satisfatório, seguindo a classificação de consistência interna apontada por Hill e Hill (2002).

Todas as questões apresentaram Coeficientes de Pearson positivos e significativos, indicando que existe a correlação entre elas, dentro de cada dimensão.

Na dimensão 1, Organização e Planejamento, a maior correlação entre os itens deu-se entre as afirmativas 1- Explicação do professor sobre os requisitos do curso e 3- Explicações do professor sobre a matéria (0,730) e a menor correlação deu-se entre os itens 1 e 4- O uso do tempo de aula pelo professor (0,557).

Na dimensão 2, Comunicação, a maior correlação entre os itens deu-se entre as afirmativas 6- A capacidade do professor de tornar clara e compreensível a aula e 7- A linguagem utilizada pelo professor para ministrar as aulas (0,693) e a menor correlação foi entre os itens 7- A linguagem utilizada pelo professor para ministrar as aulas e 9- O professor propõe perguntas ou problemas desafiadores (0,458).

Na dimensão 3, Interação professor/aluno, a maior correlação entre os itens deu-se entre as afirmativas 11- A capacidade do professor em auxiliar e responder aos alunos e 15- A disposição em ouvir as perguntas e opiniões dos alunos (0,644) e a menor correlação deu-se entre os itens 12- O respeito do professor pelos alunos e 14- A disponibilidade em atender o aluno em horário extraclasse (0,412).

Na dimensão 4, Avaliação, a maior correlação entre os itens deu-se entre as afirmativas 18- As avaliações cobram os aspectos importantes do curso e 21- As atividades aplicadas ajudam na compreensão do curso (0,592). A menor correlação deu-se entre os itens 18 e 20- A qualidade do material didático utilizado (0,445).

Na dimensão 5, Métodos e materiais didáticos, os itens de maior correlação foram a 22- Problemas ou perguntas apresentadas para discussões em grupo e a 24- Desenvolvimento de atividades em grupo (0,613). A menor correlação deu-se entre os itens 23- exercícios práticos para a compreensão dos importantes conceitos do curso e 25- uso de tecnologias para ajudar na aula.

Centra (2007) sugere que a dimensão Organização e Planejamento demonstra o quanto o professor está preparado para seus alunos no que se refere a organização do curso, sendo importante porque os discentes aprendem melhor quando percebem um curso organizado e bem planejado.



Essa dimensão permite que os alunos estabeleçam diferenças entre seus professores, conforme a habilidade de cada um (ERICKSEN; KULIK, 1974) e também permite que o professor estabeleça expectativas reais para a aprendizagem do aluno. Quando o aluno percebe uma organização por parte do professor, ele também busca se organizar melhor no decorrer do curso (CENTRA, 2007).

Centra (2007) explica que a dimensão Comunicação se refere ao professor lecionar aulas claras e compreensíveis. A utilização de exemplos é importante para a comunicação com os alunos. Exemplos anedóticos ou humorísticos levam os alunos a recordar melhor dos conceitos apresentados.

Sobre a dimensão Interação entre professores e alunos, Centra (2007) argumenta que os discentes preferem docentes que são úteis e mostram preocupação com o progresso do aluno. Os professores que estão disponíveis para dar aos alunos ajuda extra se necessário e mostram uma vontade de ouvir as perguntas dos alunos e opiniões são melhores avaliados.

Centra (2007) entende que um aspecto importante da dimensão Avaliação é a qualidade do *feedback* que os alunos recebem dos professores, também citado por Macedo (2011), Avrichir e Dewes (2006), Musetti (2007), Bittencourt et al. (2011), Freitas et al. (2014) e Silva et al. (2017).

As avaliações devem refletir o que foi visto na sala de aula e devem ser feitas de diferentes formas e com perguntas equilibradas quanto à dificuldade, de maneira a verificar o real aprendizado.

Também é importante ter critérios de avaliação claros para determinar o quão bem os testes refletem os objetivos de ensino, conforme destacam Macedo (2011), Avrichir e Dewes (2006), Centra (2007), Freitas et al. (2014) e Gomes e Borges (2008).

A metodologia usada em sala de aula é tema tratado por autores como Macedo (2011), Bittencourt et al. (2011), Avrichir e Dewes (2006), Musetti (2007), Gomes e Borges (2008), Souza e Chissolucombe (2012), Freitas et al. (2014) e Silva et al. (2017) quando referem-se ao uso de recursos didáticos/adequação pelo professor, aplicação exercícios de fixação/revisão, relacionamento entre teoria e prática e o incentivo de atividades extraclases.

Segundo Centra (2007), a escala sobre método mede as percepções da aprendizagem do aluno no curso, bem como em que medida o professor ajudou o aluno a pensar de maneira independente.

### **Conclusão e recomendações**

O objetivo desse artigo foi de realizar o processo de tradução-retrotradução do *Student Instructional Report II*, especificamente na parte em que trata da avaliação docente, buscando evidências de validação da escala. Este processo é um primeiro passo para a validação da escala e é de grande valia, já que a maioria dos instrumentos de avaliação docente voltam-se para a avaliação no ensino superior.

O processo limitou aos procedimentos de tradução-retrotradução propostos por Hill e Hill (2002). O instrumento apresentou um retorno de 5.622 casos que foram submetidos aos testes paramétricos de unidimensionalidade, confiabilidade e convergência, baseados em Hair Junior et al. (2005) e apresentados por Lima Filho e Bruni (2013).

Esses testes apresentaram a confiabilidade da escala por meio de resultados satisfatórios da análise fatorial exploratória, índice de KMO, Teste de Esfericidade de Bartlett, alfa de Cronbach e Coeficiente de Pearson.

As dimensões intituladas Organização e Planejamento, Comunicação, Interação, Avaliação e Métodos e materiais didáticos, apresentaram um único autovalor, mostrando serem unidimensionais. Os resultados do teste KMO de todas as dimensões foram superiores a 0,70 e o teste de Bartlett de todas as dimensões foram significativos indicando o uso apropriado da análise fatorial.

O alfa de Cronbach de todas as escalas foram superiores a 0,60 e considerados bem satisfatórios de acordo com os níveis sugeridos por Hill e Hill (2002). Indicando a consistência e confiabilidade das múltiplas medidas das variáveis. E o Coeficiente de Pearson de todas as escalas foram positivos e significativos, indicando a correlação entre os itens de cada escala.

Dessa forma, concluiu-se que o SIR-II utilizado para a avaliação docente possui evidências de validade quanto aos atributos de dimensionalidade, confiabilidade e convergência.

Os resultados do estudo sugerem evidências de que o SIR-II possui equivalência conceitual, semântica, alcançada por meio da discussão com os professores nas retrotraduções e aplicação do pré-teste aos alunos.

Recomenda-se que novos estudos sejam feitos de forma a complementar o processo de validação da escala, uma vez que este trabalho teve sua finalização na pré-testagem.

As limitações desse estudo referem-se à restrição do contexto de sua aplicação apenas para os alunos dos cursos integrados ao ensino médio e ao processo de tradução-retrotradução que não envolveu mais de um profissional na etapa de tradução e depois na retrotradução. Recomenda-se, portanto, que o trabalho de validação seja realizado também com alunos de cursos superiores de maneira a comparar as duas amostras independentes, ao mesmo tempo em que pelo menos mais um profissional seja envolvido na tradução-retrotradução. Espera-se que novos estudos venham também a encontrar evidências de validade dos resultados da tradução realizada.

## Referências

ANDRIOLA, C. G. *Avaliação da atuação dos docentes de Instituições de Ensino Superior (IES): o caso da Faculdade Cearense (FAC)*. 2011. 75 f. Dissertação (Mestrado em Políticas Públicas e Gestão do Ensino Superior) -Universidade Federal do Ceará, Fortaleza, 2011.

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION. *Standards for educational and psychological testing*. Washington, 2014.

AVRICHIR, I.; DEWES, F. Construção e Validação de um Instrumento de Avaliação de Desempenho Docente. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM ADMINISTRAÇÃO, 30., 2006, Brasília, DF. *Anais...* Brasília, DF: ANPAD, 2006.

BEATON, D. E. et al. Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, Philadelphia, v. 25, n. 24, p 3186-3191, 2000.

BITTENCOURT, H. R. et al. Desenvolvimento a validação de um instrumento para avaliação de disciplinas na educação superior. *Estudos em Avaliação Educacional*, São Paulo, v. 22, n. 48, p. 9-114, jan./abr. 2011.

BORSA, J. C.; DAMÁSIO, B. F.; BANDEIRA, D. R. Adaptação e validação de instrumentos psicológicos entre culturas: algumas considerações. *Paidéia*, Sorocaba, SP, v. 22, n. 53, p. 423-432, set./dez. 2012.

BOUTH, R. N. S. Avaliação docente antidemocrática: influência na prática pedagógica. *Espaço Pedagógico*, v. 20, n. 2, p. 386-390, jul/dez. 2013.

BRAGA, R.; MONTEIRO, C. A. *Planejamento estratégico sistêmico para instituições de ensino*. São Paulo: Hoper, 2005.

BRASIL. Lei nº 10.861, de 14 de abril de 2004. Institui o Sistema Nacional de Avaliação da Educação Superior – SINAES e dá outras providências. *Diário Oficial da União*. Brasília, DF, 15 abr. 2004. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2004-2006/2004/lei/l10.861.htm](http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/lei/l10.861.htm)>. Acesso em: 2 abr. 2016.

\_\_\_\_\_. Lei nº 11.892, de 29 de dezembro de 2008. Institui a Rede Federal de Educação Profissional, Científica e Tecnológica, cria os Institutos Federais de Educação, Ciência e Tecnologia, e dá outras providências. *Diário Oficial da República Federativa do Brasil*. Brasília, 30 dez. 2008.

BRANDALISE, M. A. T. Avaliação dos cursos de graduação na perspectiva dos egressos: um indicador de avaliação institucional. In: ENCONTRO NACIONAL DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM ADMINISTRAÇÃO, 9., 2012, Caxias do Sul, RS. *Anais...*, Caxias do Sul, RS: ANPED, 2012. Disponível em: <<http://www.ucs.br/etc/conferencias/index.php/anpedsul/9anpedsul/paper/download/2480/741>>. Acesso em: 21 abr. 2015.

- CASHIN, W. E. Students do rate different academic fields differently. In: THEALL, M.; FRANKLIN, J. *New directions for teaching and learning-student ratings of instruction: issues for improving practice*. San Francisco: Jossey-Bass, 1990.
- CENTRA, J. A.; GAUBATZ, N. B. Is there gender bias in student evaluations of teaching?. *The Journal of Higher Education*, Columbus, v. 71, n. 1, p. 17-33, 2000.
- CENTRA, J. A. *Enhancing your teaching through use of the SIR II report: suggestions for improvement*. Princeton: Educational Testing Service, 2007. Disponível em: <[https://www.ets.org/Media/Products/SIR\\_II/pdf/3320\\_SIRII\\_Report.pdf](https://www.ets.org/Media/Products/SIR_II/pdf/3320_SIRII_Report.pdf)> Acesso em: 10 abr. 2016.
- DANCEY, C. P.; REIDY, J. *Estatística sem matemática para psicologia*. Porto Alegre: Artmed, 2006.
- ERICKSEN, S. C.; KULK, J. A. Evaluation of teaching: memo to the faculty. *Ann Arbor: The University of Michigan*, Michigan, n. 53, 1974.
- EDUCATIONAL TESTING SERVICE. *About the SIR II Student Instructional Report: Measuring the Power of Learning*, 2017. Disponível em: <[https://www.ets.org/sir\\_ii/about/](https://www.ets.org/sir_ii/about/)>. Acesso em 03 abr. 2017.
- FETTER, R. L. et al. Adaptação transcultural para o português de instrumentos de avaliação do estado nutricional de pacientes em diálise. *Jornal Brasileiro de Nefrologia*, São Paulo, v. 36, n. 2, p. 176-185, 2014.
- FREITAS, T. A. et al. Análise de um instrumento de avaliação docente. In: CONGRESSO BRASILEIRO DE ENGENHARIA, 42., 2014, Juiz de Fora. *Anais...* Juiz de Fora: ABENGE, 2014.
- GOMES, C. M. A.; BORGES, O. Limite da validade de um instrumento de avaliação docente. *Avaliação Psicológica*, Porto Alegre, v. 7, n. 3, p. 391-401, dez. 2008.
- GUILLEMIN, F.; BOMBARDIER, C.; BEATON, D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology*, v. 46, n. 12, p. 1417-1432, 1993.
- HAIR JR., J. F. et al. *Análise multivariada de dados*. 5. ed. Porto Alegre: Bookman, 2005.
- HERDMAN, M.; FOX-RUSHBY, J.; BADIA, X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Quality of Life Research*, Dordrecht, v. 7, n. 4, p. 323-335, 1998.
- HILL, M. M.; HILL, A. *Investigação por questionário*. 2. ed. Lisboa: Edições Sílabo, 2002.
- JESUS, R. G.; ROWE, D. E. O. Comprometimento Organizacional: tradução, adaptação e validação para o contexto brasileiro da escala de sacrifícios percebidos associados com a saída. In: ENCONTRO DE ESTUDOS ORGANIZACIONAIS, 8., 2014, Gramado. *Anais...* Gramado: ANPAD, 2014.

LIMA FILHO, R. N.; BRUNI, A. L. Quanto mais faço, mais erro? uma análise sobre a presença de vieses cognitivos em julgamentos sobre orçamento. *Revista de Administração e Contabilidade da Unisinos*, São Leopoldo, RS, v. 10, n. 3, p. 224-239, 2013.

MACEDO, S. G. *Desempenho docente pela avaliação discente: uma proposta metodológica para subsidiar a gestão universitária*. 2011. 132 f. Tese (Doutorado em Engenharia da Produção) - Universidade Federal de Santa Catarina, Florianópolis, 2011.

MONTEMÓR, M. L. *Autoavaliação institucional: dimensões políticas e gestão democrática*. 2013. 121 f. Dissertação (Mestrado em Educação) - Universidade Federal de São Carlos, São Carlos, 2013.

MOREIRA, M. A. Avaliação do professor pelo aluno como instrumento de melhoria do ensino universitário. *Educação e Seleção*, São Paulo, v. 4, n. 119, p. 109-124, jul./dez., 1981.

MUSETTI, H. *Evidências de validade de um instrumento de avaliação docente*. 2007. 86 f. Dissertação (Mestrado em Psicologia) – Universidade São Francisco, Itatiba, SP, 2007.

OLIVEIRA, A. F.; HILDENBRAND, L.; LUCENA, R. S. Adaptação transcultural de instrumentos de medida e avaliação em saúde: estudo de metodologias. *Revista Acreditação*, Rio de Janeiro, v. 5, n. 10, p. 13-33, 2015.

SERRANO, L. E.; VIZCARRA, V. A. Variables asociadas a la competencia docente universitaria en la opinión de los estudiantes. *Archivos Analíticos de Políticas Educativas*, México, v. 2, n. 1, p. 1-26, jan. 2013.

SILVA, M. A. et al. Construção e estudos de evidências de validade da escala de avaliação docente. *Revista Brasileira de Educação*, Rio de Janeiro, v. 22, n. 70, p. 690-707, jul./set. 2017.

SOUSA, J. V. de. Avaliação institucional, estratégias de marketing e imagens projetadas pelas IES: que relação é esta?. In: OLIVEIRA, J. F.; FONSECA, M. (Org). *Avaliação institucional: sinais e práticas*. São Paulo: Xamã, 2008.

SOUZA, R. R.; CHISSOLUCOMBE, I. Instrumentos de avaliação docente: avaliação pelo discente, coordenador de curso e auto-avaliação. *Revista Projeção e Docência*, Brasília, DF, v. 3, n. 1, p. 12-17, mar. 2012.

SWAINE-VERDIER, A. et al. Adapting quality of life instruments. *Value in Health*, [S.l.], v. 7, supl. 1, p. 27-30, 2004.

TEJEDOR, F.J.; MONTEIRO, M.L. Indicadores de la calidad docente para la evaluación del profesor universitario. *Revista española de pedagogía*, Del Valle, n. 186, p. 259-27, 1990.

Recebido em: 19/04/2018

Aceito para publicação em: 14/02/2019

## Teacher Evaluation by the Student in the Integrated Technical Education: evidences of the SIR-II scale validity

### Abstract

The aim of this article was to do the translation, back-translation, committee review and pre-testing of the second version of the Educational Testing Service (ETS) Student Instructional Report (SIR-II) Scale, specifically in the teacher evaluation part. The instrument analyzed was composed of five dimensions: Organization and Planning; Communication; Teacher / Student Interaction; Assessments, Examinations and Grades; and Supplementary Teaching Methods. The sample was composed by 5,622 cases generated by students of technical courses integrated to the high school from various campuses of the Federal Institute of Education, Science and Technology of the Northern Minas Gerais (IFNMG), a public educational institution. The model presented an adjustment for unidimensionality, reliability and convergence, presenting evidences of scale validity in its Brazilian version.

**Keywords:** Evaluation. Scale. Translation. Validation.

## Evaluación del Docente por el Discente en el Ámbito de la Enseñanza Técnica Integrada: evidencias de validación de la escala SIR-II

### Resumen

Este artículo tuvo el objetivo de hacer la traducción, la traducción inversa, revisión por expertos y la pre-prueba de la segunda versión de la *Student Instructional Report* (SIR-II) de la organización *Educational Testing Service* (ETS), especialmente en lo referente a la evaluación del docente. El instrumento analizado se compuso de cinco dimensiones: Organización y Planificación; Comunicación; Interacción Profesor / Alumno; Evaluaciones, Exámenes y Notas; y Métodos Complementarios de Enseñanza. La muestra del presente estudio fue compuesta por 5.622 casos generados por estudiantes de cursos técnicos integrados a la enseñanza media de diferentes campi del Instituto Federal de Educación, Ciencia y Tecnología del Norte de Minas Gerais (IFNMG). El modelo presentó ajuste relativo a la unidimensionalidad, confiabilidad y convergencia, presentando evidencias de validez de la escala en su versión brasileña.

**Palabras clave:** Evaluación. Escala. Traducción. Validación.