

# AVALIAÇÃO NO ENSINO SUPERIOR: CONCEPÇÕES MÚLTIPLAS DE ESTUDANTES BRASILEIROS<sup>1</sup>

DANIEL ABUD SEABRA MATOS  
SÉRGIO DIAS CIRINO  
GAVIN THOMAS LUMSDEN BROWN  
WALTER LANA LEITE

## RESUMO

*Os objetivos dessa pesquisa foram: 1) analisar as concepções de avaliação dos alunos do ensino superior; 2) adaptar e validar o questionário Students' Conceptions of Assessment (SCoA); 3) investigar as definições de avaliação dos alunos; 4) analisar como as concepções de avaliação predizem as definições de avaliação dos alunos; 5) analisar diferenças nas concepções de avaliação dos alunos pertencentes a uma Instituição de Ensino Superior pública e a uma privada. Utilizamos a teoria da aprendizagem autorregulada e métodos quantitativos e qualitativos.*

**1** Apoio: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

**PALAVRAS-CHAVE** CONCEPÇÕES DE AVALIAÇÃO • AVALIAÇÃO • ENSINO SUPERIOR.

## **RESUMEN**

Los objetivos de esta investigación fueron: 1) analizar las concepciones de evaluación de alumnos de Enseñanza Superior; 2) adaptar y validar el cuestionario *Students' Conceptions of Assessment (SCoA)*; 3) investigar las definiciones de evaluación de los alumnos; 4) analizar cómo las concepciones de evaluación predicen las definiciones de evaluación de los alumnos; 5) analizar diferencias en las concepciones sobre evaluación de alumnos que pertenecen a una Institución de Enseñanza Superior pública y a otra privada. Utilizamos la teoría de autorregulación del aprendizaje y métodos cuantitativos y cualitativos.

**PALABRAS CLAVE** CONCEPCIONES DE LA EVALUACIÓN •  
EVALUACIÓN • EDUCACIÓN SUPERIOR.

## **ABSTRACT**

The goals of this study were to 1) overview Brazilian university students' conceptions of assessment; 2) adapt and validate the *Students' Conceptions of Assessment (SCoA)* inventory; 3) establish the underlying dimensions shaping students' definitions of assessment; 4) analyze how conceptions of assessment predicted students' definitions of assessment; and 5) investigate differences between students at public and private universities. We used the theory of self-regulated learning and quantitative and qualitative methods.

**KEYWORDS** CONCEPTIONS OF ASSESSMENT • ASSESSMENT •  
HIGHER EDUCATION.

## **1 INTRODUÇÃO**

Dentre as várias possibilidades de abordagem da avaliação, a análise das concepções de avaliação dos estudantes tem sido uma das opções de trabalho na educação. Entretanto, existem poucas pesquisas que consideram a avaliação sob o ponto de vista dos próprios alunos. As pesquisas sobre as concepções de avaliação dos alunos são escassas (BROWN; HIRSCHFELD, 2005).

O termo concepção é usado aqui para se referir a “sistemas complexos de explicação”, uma “rede complexa de ideias, conceitos, representações e preconceitos”. Dessa forma, as concepções podem informar a maneira como as pessoas percebem, avaliam e agem com relação a um determinado fenômeno. As concepções de avaliação dos alunos são de grande importância porque a avaliação tem um papel essencial na qualidade do aprendizado, e as crenças e atitudes dos alunos com relação aos propósitos da avaliação parecem influenciar suas práticas e resultados (ENTWISTLE; ENTWISTLE, 1991). Assim, diversas pesquisas descrevem que as concepções dos alunos sobre a avaliação influenciam de maneira significativa a abordagem deles com relação ao aprendizado (STRUYVEN et

al., 2005). Associações entre as concepções de avaliação e o desempenho acadêmico dos alunos também aparecem na literatura (BROWN; HIRSCHFELD, 2007, 2008).

Considerando especificamente o ensino superior brasileiro, diversas reformas e mudanças vêm acontecendo na área da avaliação: novas legislações, maior uso de avaliação formativa, novas práticas avaliativas (exemplos: portfólio, avaliação feita pelos colegas, autoavaliação), novos formatos de cursos e de Instituições de Ensino Superior (IES), novas exigências do mercado, entre outras. A avaliação institucional também tem sido cada vez mais incorporada à realidade das IES e os alunos têm o Enade como componente curricular obrigatório. No entanto, a eficácia de práticas avaliativas tidas como inovadoras está condicionada à forma como os alunos entendem e respondem à avaliação. Por isso, investigar como os estudantes do ensino superior concebem e experienciam a avaliação se tornou relevante.

Em face da importância dessa área de estudos, realizamos uma revisão da literatura sobre as concepções de avaliação dos alunos universitários brasileiros e também uma pesquisa empírica cujos objetivos eram: 1) analisar as concepções de avaliação dos alunos do ensino superior; 2) adaptar e validar o questionário *Students' Conceptions of Assessment (SCoA), version VI* (Anexo A) (BROWN, 2006) para a realidade brasileira; 3) investigar as definições de avaliação dos alunos do ensino superior; 4) analisar como as concepções de avaliação predizem as definições de avaliação dos alunos; 5) analisar diferenças nas concepções de avaliação dos alunos pertencentes a uma IES pública e a uma IES privada.

### 1.1 TEORIA SOBRE AS CONCEPÇÕES DE AVALIAÇÃO

Encontramos na literatura sobre as concepções de avaliação dos alunos quatro grandes grupos de concepções: 1) a avaliação melhora o ensino e a aprendizagem; 2) a avaliação tem um impacto emocionalmente positivo nos alunos; 3) a avaliação é irrelevante e os alunos respondem negativamente a ela; 4) a avaliação torna os alunos responsáveis (*accountability*) (BROWN, 2008; BROWN; HIRSCHFELD, 2008). Os alunos demonstram ter consciência de todas essas concepções e tendem a se orientar por uma ou pela combinação de várias delas. Pesquisas sobre o ensino fundamental em diferentes países (ATKINSON, 2003;

REAY; WILLIAM, 1999) e trabalhos sobre o ensino médio e o ensino superior revelaram concepções múltiplas e conflitantes de avaliação (CAMARGO, 1997; MONI et al., 2002).

A concepção de melhora enfatiza a expectativa dos alunos de que a avaliação leva a uma aprendizagem melhor e mais eficaz (BROWN, 2008). A concepção de que a avaliação tem um impacto emocionalmente positivo nos estudantes enfatiza as respostas emocionais dos alunos com relação a formatos diferentes de avaliação (BROWN et al., 2009a). Respostas emocionais positivas, como a motivação dos colegas de sala de aula e a avaliação como uma experiência agradável, estão incluídas nessa concepção. A concepção de irrelevância enfatiza a percepção de que a avaliação é negativa, ruim ou injusta. Existem evidências de que os estudantes diferem quanto ao grau de percepção negativa da avaliação. A maioria percebe a avaliação de forma cada vez mais negativa na medida em que avança no sistema educacional (BROWN et al., 2009a). A concepção de responsabilização enfatiza a ideia de que a avaliação torna os alunos responsáveis por seu aprendizado, em consequência de notas ou certificados (BROWN, 2008). Além disso, nessa concepção, a avaliação também pode estar relacionada com fatores externos ao indivíduo, como a escola, os pais ou o futuro profissional. (BROWN et al., 2009b). A existência dessa concepção pode ser inferida das respostas que os alunos dão para avaliações que têm consequências importantes para eles.

Os alunos podem se orientar por uma concepção ou por uma combinação de diversas concepções – como já foi dito. Assim, um mesmo estudante pode relatar que a avaliação serve para melhorar o aprendizado e que é injusta (concepções de avaliação à primeira vista conflitantes). A natureza complexa da educação parece contribuir para isso.

## 1.2 O QUESTIONÁRIO *STUDENTS' CONCEPTIONS OF ASSESSMENT (SCoA)*

O SCoA (Anexo A) foi desenvolvido na Nova Zelândia para avaliar as concepções dos alunos sobre a avaliação. O formato de resposta do questionário é uma escala *Likert* de seis pontos: discordo fortemente, discordo na maior parte, concordo ligeiramente, concordo moderadamente, concordo na maior parte, concordo fortemente. O instrumento possui 33 itens

distribuídos em oito fatores: Melhora-aluno (itens 1, 10, 14, 15, 19); Melhora-professor (itens 5, 8, 9, 23, 27, 30); Emoção-aluno (itens 6, 31); Emoção-sala (itens 2, 12, 17, 21, 25, 28); Irrelevante-ruim (itens 3, 13, 18, 22, 26); Irrelevante-ignorar (itens 7, 29, 32); Responsabilização-aluno (itens 4, 16, 20, 33); Responsabilização-escola (itens 11, 24). Assim, as escalas do SCoA-VI avaliam as concepções de avaliação presentes na literatura educacional, tornando cada uma delas mais específica por meio de subdivisões (durante o processo de adaptação e validação na amostra brasileira, o item 33 foi excluído).

O SCoA, além de medir concepções de avaliação, mede definições de avaliação. A seguir, explicamos de forma mais detalhada a diferença entre concepções e definições de avaliação.

As concepções de avaliação (ou seja, seus propósitos ou intenções) dependem do tipo de avaliação que é considerado. Por exemplo: se avaliação significa teste, em uma situação, ou significa interação com o professor, em outra, então, provavelmente, um conjunto diferente de propósitos da avaliação pode vir à nossa mente.

Uma maneira de controlar esse efeito potencial é estabelecer o que a palavra avaliação significa para os participantes. Assim, além dos itens específicos sobre as concepções de avaliação, sobre a natureza e o propósito da avaliação (itens 1 ao 32), uma lista de práticas avaliativas pode ser mostrada aos sujeitos. O SCoA faz isso por meio da “Parte Complementar do Questionário”. A pergunta dessa parte é: “Quando você pensa na palavra avaliação, quais tipos de atividades avaliativas vêm à sua mente?”. Logo em seguida, são apresentados 12 tipos de atividades avaliativas. Os alunos podem marcar quantas opções quiserem (exemplos: trabalho em grupo, autoavaliação). Portanto, os alunos indicam suas definições de avaliação apontando as práticas avaliativas que associam com o termo avaliação (ver seção 3.1.1).

A lógica por trás desse procedimento é que entender como os alunos definem avaliação (tipos de atividades avaliativas) ajuda no esclarecimento das concepções de avaliação. Diversos estudos já utilizaram essa abordagem com professores (BROWN, 2002) e estudantes do ensino médio (BROWN et al., 2009a).

O desenvolvimento do SCoA-VI (BROWN, 2006) foi baseado na teoria da aprendizagem autorregulada (*self-regulated learning*).

Em síntese, a teoria da aprendizagem autorregulada destaca a importância da iniciativa pessoal na aprendizagem. A aprendizagem autorregulada analisa como os estudantes selecionam, organizam ou criam ambientes de aprendizagem vantajosos para eles mesmos (ZIMMERMAN, 1990).

Os resultados do questionário SCoA, portanto, são interpretados a partir dessa teoria. Pesquisas anteriores apontam que alunos que veem a avaliação como uma força construtiva para a responsabilidade pessoal obtêm melhores notas. Já aqueles que procuram “culpar” escolas ou professores pelos resultados avaliativos, aqueles que não assumem a avaliação seriamente ou que ignoram a avaliação recebem notas mais baixas (BROWN; HIRSCHFELD, 2005). Deve-se destacar que não temos a intenção de estabelecer uma relação de causalidade. O que as pesquisas anteriores demonstram é que existe uma associação entre as concepções de avaliação e o desempenho acadêmico dos alunos. Assim, é difícil dizer se os baixos resultados avaliativos resultam de concepções inapropriadas ou se os baixos resultados causam concepções inapropriadas.

Brown, Peterson e Irving (2009b) discutem a hipótese de que somente a concepção de melhora parece estar alinhada com uma resposta positiva de autorregulação do aprendizado. Já as concepções de que a avaliação tem um impacto emocionalmente positivo nos alunos, de irrelevância e de responsabilização, seriam potencialmente concepções com uma relação negativa com a autorregulação do aprendizado.

Por fim, alguns itens do SCoA levam em consideração a função formativa da avaliação. Exemplo: “Os professores usam meus resultados da avaliação para ver o que precisam me ensinar em seguida”.

## **2 MÉTODO**

### **2.1 O DESENHO DE PESQUISA**

A pesquisa teve um delineamento de *survey*. Esse tipo de delineamento envolve procedimentos em pesquisa quantitativa na qual o pesquisador aplica um *survey* em uma amostra ou em uma população inteira para descrever atitudes, opiniões, comportamentos ou características da população. Após a coleta de dados,

o pesquisador realiza análises estatísticas para descrever tendências nas respostas das questões e testar problemas de pesquisa ou hipóteses. Ele também interpreta o significado dos dados relacionando os resultados dos testes estatísticos com estudos anteriores (CRESWELL, 2005).

Embora existam muitas formas diferentes de *survey*, os pesquisadores geralmente coletam dados usando duas formas básicas: questionários e entrevistas. No presente trabalho, fizemos uso de um questionário para a coleta de dados. Optamos por uma aplicação em grupo autoadministrada (DILLMAN, 2000). Ou seja: apesar do instrumento ser autoadministrado (o próprio aluno respondia), o questionário foi aplicado em grupo e presencialmente nas respectivas salas de aula dos estudantes.

## 2.2 CARACTERIZAÇÃO DA AMOSTRA

A aplicação do SCoA brasileiro foi realizada durante o segundo semestre de 2008. Estudantes do ensino superior ( $N = 756$ ; 216 homens e 540 mulheres; idade  $M = 24,53$  anos;  $DP = 5,91$ ) de duas IES de Minas Gerais participaram: uma Universidade Federal ( $N = 297$ ) e um Centro Universitário privado ( $N = 459$ ). Dezoito cursos fizeram parte da amostra, incluindo alguns de graduação tecnológica: Administração de Empresas, Arquitetura e Urbanismo, Ciências Biológicas, Ciências Contábeis, Ciências Econômicas, Comércio Exterior, Engenharia Civil, Farmácia, Física, Gestão Comercial, Gestão de Recursos Humanos, Medicina Veterinária, Nutrição, Pedagogia, Processos Gerenciais, Psicologia, Serviço Social e Terapia Ocupacional.

## 2.3 ETAPAS DA PESQUISA

Inicialmente, fizemos uma revisão da literatura sobre as concepções de avaliação dos alunos universitários brasileiros. Posteriormente, um conjunto de três estudos foi realizado usando métodos quantitativos (Escalonamento Multidimensional, Análise Fatorial Exploratória e Confirmatória, Análise de Variância, Modelagem de Equações Estruturais) e qualitativos (análise de conteúdo para investigar as questões abertas do SCoA). Por uma limitação de espaço, apresentaremos apenas os resultados quantitativos.



## 2.4 PROCEDIMENTOS TEÓRICOS

Usamos o modelo de validação de instrumentos proposto por Pasquali (1999) que se baseia em três grandes polos: procedimentos teóricos, procedimentos empíricos (experimentais) e procedimentos analíticos (estatísticos).

Inicialmente, fizemos uma tradução da versão em inglês do SCoA para o português envolvendo todos os passos necessários, como o procedimento de tradução inversa (*backtranslation*). Participou dessa tradução uma equipe de pesquisadores.

Posteriormente, realizamos uma análise teórica dos itens. Doze alunos participaram de entrevistas semiestruturadas em grupos: cinco alunos da IES pública e sete alunos da IES privada. Após responderem ao SCoA, eles foram instruídos a marcar qualquer dúvida. Depois, todas as dúvidas foram discutidas em grupo. Utilizando essas entrevistas, pequenas alterações foram realizadas no SCoA.

## 2.5 PROCEDIMENTOS EMPÍRICOS

Realizamos um estudo piloto em uma sala da IES pública com 37 alunos e em uma sala da IES privada com 33 alunos. Após o piloto, não foram necessárias mudanças adicionais no instrumento. As entrevistas em grupo e o estudo piloto foram feitos antes de usar o SCoA em uma amostra grande de alunos.

Depois, foi feito contato com as IES solicitando a aprovação da pesquisa. A aplicação ocorreu de forma coletiva na sala dos estudantes. O SCoA foi aplicado pessoalmente pelo pesquisador e por uma colaboradora. Durante a aplicação, obtivemos a autorização dos alunos para serem incluídos na pesquisa por meio da assinatura do termo de consentimento livre e esclarecido.

Os procedimentos analíticos são abordados a seguir.

# 3 RESULTADOS

## 3.1 DEFINIÇÕES DE AVALIAÇÃO

### 3.1.1 ESCALONAMENTO MULTIDIMENSIONAL

Como parte integrante do SCoA, os alunos tinham de responder à pergunta: “Quando você pensa na palavra avaliação, quais tipos de atividades avaliativas vêm à sua mente?”.

As respostas foram consideradas de forma dicotômica (0 = não selecionada; 1 = selecionada). A Tabela 1 mostra a frequência e a porcentagem das atividades avaliativas.

**TABELA 1 – Tipos de atividades avaliativas e quadrantes do Escalonamento Multidimensional**

TIPOS DE ATIVIDADES AVALIATIVAS	FREQUÊNCIA	PORCENTAGEM
<b>Formal/controle do professor</b>		
*O professor aplica prova com questões abertas	610	87
*O professor atribui uma nota ou avalia um trabalho em grupo	542	77
O professor atribui uma nota ou avalia um trabalho escrito que eu entrego	533	76
*O professor aplica prova com questões de múltipla escolha	525	75
O professor me dá uma nota em um trabalho feito na sala de aula	448	64
Média	531.6	75.8
<b>Informal/controle do professor</b>		
O professor me observa na sala de aula e julga o meu aprendizado	260	37
Eu atribuo uma nota ou avalio meu próprio desempenho (autoavaliação)	227	32
O professor atribui uma nota ao meu desempenho depois de se reunir ou discutir comigo sobre o meu trabalho	210	30
Meus colegas atribuem uma nota ou avaliam meu desempenho	92	13
Média	197.2	28
<b>Formal/controle do aluno</b>		
*O professor aplica prova em dupla	321	46
*O professor aplica prova com consulta	318	45
Média	319.5	45.5
<b>Informal/controle do aluno</b>		
O professor me faz perguntas em voz alta na sala de aula	187	27
Média	187	27

Nota: os itens com \* foram criados para a versão brasileira.

Fonte: elaboração dos autores.

Um procedimento exploratório de Escalonamento Multidimensional (Kruskal stress = .08;  $R^2 = .97$ ) foi usado para determinar as dimensões subjacentes para as respostas dos alunos aos 12 tipos de atividades avaliativas. O *software* SPSS versão 17.0

foi utilizado. O presente estudo indicou que duas dimensões (formalidade e lócus de controle) e sua interação poderiam explicar como os estudantes definem a avaliação, resultando em quatro grupos de atividades avaliativas: informal/controle do aluno; informal/controle do professor; formal/controle do aluno; formal/controle do professor (Tabela 1). A avaliação foi definida pelos alunos principalmente como métodos formais/convencionais controlados pelo professor do tipo exames e testes. Até mesmo práticas supostamente controladas pelos alunos (autoavaliação, avaliação dos colegas) foram percebidas como práticas controladas pelo professor.

### 3.1.2 ANÁLISE FATORIAL CONFIRMATÓRIA

Utilizamos prioritariamente a Análise Fatorial Confirmatória (AFC) na validação do SCoA. É um procedimento mais robusto do que técnicas como a consistência interna para avaliar a validade de um instrumento porque usa os padrões dos fatores, das correlações, das covariâncias e os valores de resíduo ou erro dentro de uma matriz de dados para determinar como um modelo se ajusta aos dados (KLEM, 2000; THOMPSON, 2000).

Os índices de ajuste do modelo utilizados para a Análise Fatorial e para a Modelagem de Equações Estruturais foram os mesmos: a estatística qui-quadrado, o índice de ajuste comparativo (CFI), o índice de Tucker-Lewis (TLI), o índice *gammahat*, a raiz do erro quadrático médio de aproximação (RMSEA) e a raiz do resíduo quadrático médio padronizado (SRMR) (FAN, SIVO, 2007).

O critério de corte de  $CFI \geq .90$ ,  $TLI \geq .90$ ,  $gammahat \geq .90$ ,  $RMSEA \leq .08$  e  $SRMR \leq .08$  foi usado no presente trabalho para estabelecer um ajuste adequado ao invés do critério de corte proposto por Hu e Bentler (1999):  $TLI \geq .95$ ,  $CFI \geq .95$ ,  $RMSEA \leq .05$  e  $SRMR \leq .08$ . Esse critério foi largamente utilizado durante alguns anos. Porém, recentemente, foi demonstrado que o critério de Hu e Bentler é muito restritivo quando aplicado a instrumentos multifatoriais (*multi-factor rating instruments*) (MARSH et al., 2004, 2005). Em outras palavras, o critério de Hu e Bentler penaliza exageradamente modelos complexos (como é o caso da estrutura fatorial do SCoA).

Destacamos, no entanto, que ainda existe muito debate sobre os padrões mais apropriados para o ajuste de modelos. Essa questão é ainda mais controversa no caso de modelos

hierárquicos multifatoriais com amostras grandes. A estatística qui-quadrado é sensível a erros muito pequenos de especificação do modelo se o tamanho da amostra for grande (BARRETT, 2007). Os índices de *goodness-of-fit* CFI e TLI são muito sensíveis a modelos complexos (mais de três fatores ou com estrutura hierárquica) (CHEUNG; RENSVD, 2002). Já os índices RMSEA, SRMR (*badness-of-fit* índices) e *gamma hat* (*goodness-of-fit* índice) têm se mostrado mais resistentes ao impacto de amostras grandes, modelos complexos e erros de especificação do modelo. Fan e Sivo (2007) recomendam o índice *gamma hat* como o mais resistente a erros de especificação do modelo. Valores do *gamma hat* > .95 indicam um ajuste excelente e valores > .90 são bons. A fórmula desse índice de *goodness-of-fit* utiliza o RMSEA, os graus de liberdade e o número de variáveis observadas. Por levar em consideração o número de variáveis do modelo, esse índice penaliza menos os modelos complexos.

O *software* MPLUS versão 3.11 foi utilizado para realizar a AFC. O método de Máxima Verossimilhança, que é o estimador padrão do MPLUS, assume uma normalidade multivariada. No entanto, esse pressuposto é violado com o SCoA, que contém itens do tipo Likert de seis pontos e variáveis dicotômicas em sua parte complementar. Então, a AFC foi conduzida com um estimador robusto do MPLUS que usa correlação policórica para a análise de indicadores categóricos, denominado *Mean and variance-adjusted weighted least squares estimator* (WLSMV).

Ainda com relação às definições de avaliação dos alunos (dimensões subjacentes nas respostas aos 12 tipos de atividades avaliativas), a solução com quatro fatores encontrada no Escalonamento Multidimensional foi inadmissível (não convergiu) na AFC. Um modelo com três fatores foi o que apresentou o melhor ajuste:  $\chi^2(34, N= 756) = 176.762$ ,  $p < .0000$ , CFI = .92, TLI = .92, *Gamma hat* = .97, RMSEA = .075, SRMR = .100. A estatística qui-quadrado do modelo de três fatores foi significativa. Entretanto, como já discutimos, a estatística qui-quadrado é sensível a erros muito pequenos de especificação do modelo se o tamanho da amostra for grande. Quatro índices sugerem um bom ajuste do modelo (CFI = .92; TLI = .92; *gamma hat* = .97; RMSEA = .075). Juntos esses índices sugerem que as respostas ao SCoA brasileiro se ajustam bem à estrutura fatorial de três fatores: informal/controle do professor; formal/controle do aluno; formal/

controle do professor. Para dados adicionais sobre a validade do instrumento, como cargas fatoriais padronizadas dos itens, consistência interna e validade discriminante, ver Matos (2010).

### 3.2 CONCEPÇÕES DE AVALIAÇÃO

#### 3.2.1 ANÁLISE FATORIAL CONFIRMATÓRIA

Com relação às concepções de avaliação dos alunos (itens 1 ao 32), analisamos os índices de ajuste por meio de vários modelos. A AFC foi conduzida usando o estimador robusto do MPLUS WLSMV, que fornece uma estimação sem viés para escalas Likert de diferentes números de categorias e graus de simetria e curtose (FLORA; CURRAN, 2004).

Um modelo com oito fatores teve o melhor ajuste:  $\chi^2(138, N=756) = 1116.312, p < .0000, CFI = .83, TLI = .95, Gammahat = .92, RMSEA = .097, SRMR = .057$ . A estatística qui-quadrado do ajuste do modelo foi significativa. Três índices sugerem um bom ajuste do modelo (TLI = .95; gamma hat = .92; SRMR = .057). Juntos esses índices sugerem que as respostas ao SCoA brasileiro se ajustam bem à estrutura fatorial proposta de oito fatores.

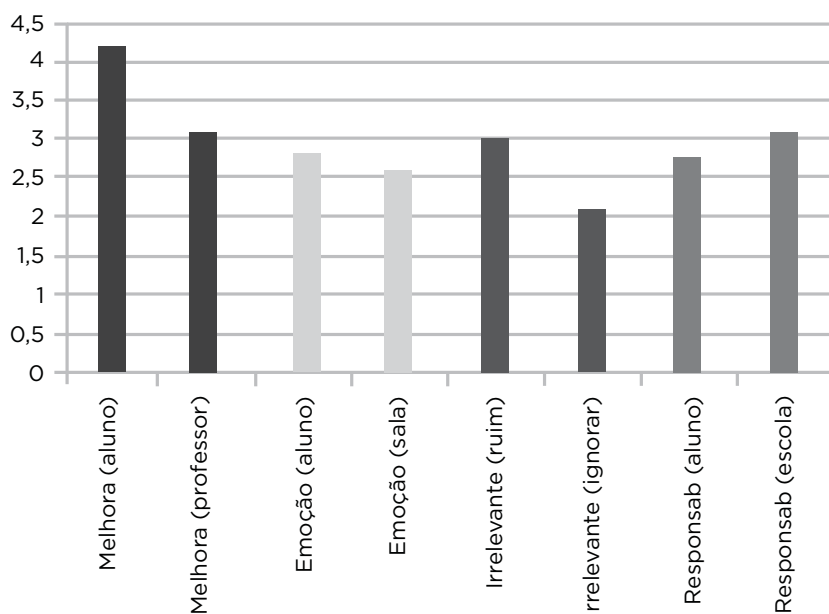
Para dados adicionais sobre a validade do instrumento, como cargas fatoriais padronizadas dos itens, consistência interna, validade discriminante e análise fatorial exploratória, ver Matos (2010).

#### 3.2.2 ANÁLISES ADICIONAIS

As médias das respostas para cada fator foram examinadas (Figura 1). O eixo vertical representa o nível de concordância com a concepção (1 = discordo fortemente; 2 = discordo na maior parte; 3 = concordo ligeiramente; 4 = concordo moderadamente; 5 = concordo na maior parte; 6 = concordo fortemente).

Como mostra a Figura 1, os alunos parecem conceber que a avaliação, em primeiro lugar, melhora o ensino e a aprendizagem (Melhora) e, em segundo lugar, torna os alunos e escolas responsáveis (Responsabilização). Quanto à concepção de que a avaliação é irrelevante, o fator “Irrelevante (ruim)” apresenta como média aproximadamente a resposta “concordo ligeiramente”; e o fator “Irrelevante (ignorar)”, a resposta “discordo na maior parte”. Já a concepção “Emoção” parece estar associada aproximadamente com um valor de ligeira concordância.

FIGURA 1 - Médias das respostas das concepções de avaliação



Nota: N=756.

Fonte: elaboração dos autores.

Em outras palavras, os alunos parecem estar pensando algo como: “a avaliação ajuda a melhorar o ensino-aprendizagem, torna alunos e escolas responsáveis e não é algo a ser ignorado. Mas a avaliação também pode ser ruim ou injusta e quase não tem um impacto emocionalmente positivo nos alunos”.

### 3.2.3 DIFERENÇAS ENTRE A IES PÚBLICA E A IES PRIVADA

Para comparar as concepções de avaliação de alunos pertencentes a uma IES pública e uma IES privada, realizamos uma Análise de Variância (Anova). Os alunos da IES privada acreditam mais do que os alunos da IES pública que a avaliação melhora o ensino-aprendizagem e torna alunos e escolas responsáveis. Além disso, os alunos da instituição pública acreditam em um maior grau que a avaliação é algo ruim ou injusto. Por fim, os alunos da IES privada percebem uma maior associação da avaliação a benefícios como a motivação dos colegas e o clima da turma. Os alunos da instituição privada também concebem a avaliação como uma experiência mais agradável quando comparados a alunos da instituição pública.

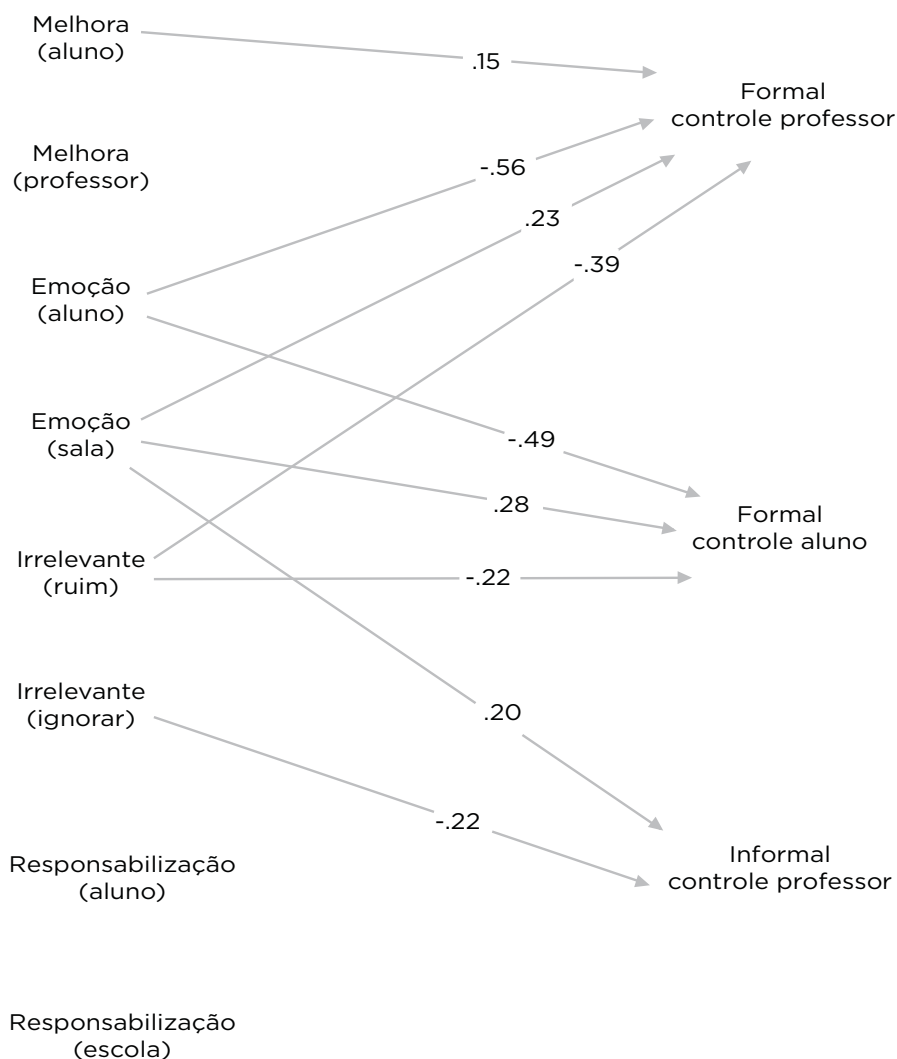
Apesar de a presente pesquisa não ter coletado medidas cognitivas dos estudantes da amostra, algumas considerações

podem ser feitas. No Brasil, o desempenho cognitivo médio dos alunos das IES públicas é geralmente melhor do que o desempenho cognitivo médio dos alunos das IES privadas. Assim, embora pudéssemos esperar que um impacto emocionalmente positivo nos alunos (individualmente ou na sala de aula) elevasse o desempenho acadêmico, parece que essa concepção de avaliação é desadaptativa. As concepções “Emoção-aluno” e “Emoção-sala” estão associadas a um desempenho acadêmico mais baixo em leitura e matemática (BROWN; HIRSCHFELD, 2007, 2008). Já com relação à ligação entre as concepções de avaliação e a teoria da aprendizagem autorregulada, Brown, Peterson e Irving (2009b) argumentam que somente a concepção de melhora parece estar alinhada com uma resposta positiva de autorregulação do aprendizado. Dessa forma, apesar de os alunos da IES privada acreditarem mais do que os alunos da IES pública na concepção “Melhora”, eles também acreditam mais nas concepções “Emoção” e “Responsabilização” (que seriam potencialmente concepções com uma relação negativa com a autorregulação do aprendizado).

### **3.3 RELAÇÃO ENTRE AS CONCEPÇÕES E AS DEFINIÇÕES DE AVALIAÇÃO**

Analisamos também como as concepções de avaliação predizem as definições de avaliação dos alunos. Vários modelos foram testados. O modelo apresentado na Figura 2 foi o que apresentou o melhor ajuste:  $\chi^2(205, N = 756) = 940.095, p < .0000, CFI = .90, TLI = .96, \text{Gamma hat} = .96, RMSEA = .069, SRMR = .062$ . Cinco índices indicam um bom ajuste, sugerindo que as relações estruturais propostas se ajustam bem ao modelo. A Modelagem de Equações Estruturais foi conduzida usando o estimador robusto do MPLUS WLSMV.

**FIGURA 2 – Modelagem de Equações Estruturais**



Nota: Todos os coeficientes de regressão apresentados foram estatisticamente significantes,  $p < 0.05$ .

Fonte: elaboração dos autores.

Como indicado na Figura 2, a definição de avaliação como práticas formais (controle do aluno e professor) diminuiu com uma concordância de que a avaliação era uma experiência agradável (Emoção-aluno) e algo ruim (Irrelevante-ruim). Em outras palavras, é como se os estudantes estivessem dizendo: “Se a avaliação é formal, eu não necessariamente penso que seja uma experiência agradável para mim. Apesar disso, eu também não considero as avaliações formais como algo ruim”.

A definição “Formal/controlado do professor” aumenta ainda com uma concordância de que a avaliação era para a melhora do aluno (Melhora-aluno). Assim, poderíamos acrescentar a



seguinte parte na fala hipotética do aluno no parágrafo anterior: “Além disso, se a atividade avaliativa é formal e controlada pelo professor, ela me ajuda a melhorar meu aprendizado, me ajuda nos meus próximos passos de aprendizagem”. Merece destaque aqui o fato de que a concepção “Melhora” só é associada no pensamento dos alunos com a definição de avaliação como uma prática formal e controlada pelo professor.

Todos os tipos de definição de avaliação (formais e informais) foram preditos positivamente por um aumento percebido em benefícios da sala de aula (Emoção-sala). Dessa forma, os alunos parecem perceber que todos os tipos de atividades avaliativas estão associados com benefícios como a motivação dos colegas e o clima da turma.

Finalmente, a definição de avaliação como práticas informais controladas pelo professor diminuiu com uma concordância de que a avaliação era algo a ser ignorado (Irrelevante-ignorar). Em outras palavras, é como se os alunos estivessem dizendo: “Se a avaliação é informal e controlada pelo professor, ela é boa para a dinâmica e o clima da sala de aula. Além disso, eu não considero esse tipo de atividade avaliativa como algo a ser ignorado”. Essa relação entre a definição “informal” e a concepção “Irrelevante-ignorar” foi no sentido oposto dos resultados encontrados na Nova Zelândia para estudantes do ensino médio, onde os alunos consideraram esse tipo de atividade avaliativa como algo a ser ignorado (BROWN et al., 2009a).

#### **4 CONSIDERAÇÕES FINAIS**

A inclusão do SCoA em pesquisas que investigam o desempenho cognitivo dos alunos é uma alternativa possível. Além disso, instrumentos que avaliam as concepções de avaliação podem contribuir para a reflexão da prática pedagógica e serem inseridos em programas de formação de professores.

Pesquisas adicionais são necessárias para investigar de forma mais aprofundada questões como: as diferenças entre IES públicas e privadas e as definições de avaliação (incluindo novos tipos de atividades avaliativas).

Nosso trabalho contribuiu para suprir uma lacuna existente nas pesquisas educacionais brasileiras: o pouco uso de

metodologias quantitativas, principalmente de métodos de análise de dados mais robustos. Outra contribuição dessa pesquisa foi mostrar que os aspectos psicológicos, a compreensão dos alunos sobre a avaliação, são mais gerais e universais do que pensado anteriormente.

Concluimos apontando uma questão central para os educadores hoje: como os estudantes estão percebendo a avaliação no contexto atual? Se não conseguirmos responder a essa questão, algumas reformas avaliativas podem não ter o resultado esperado. Como já discutimos, a eficácia das práticas avaliativas está condicionada à forma como os estudantes entendem e, conseqüentemente, respondem à avaliação. Isso fica evidente por meio de alguns resultados da nossa pesquisa. Exemplo: nosso modelo demonstra que, no pensamento dos estudantes, atividade formal e controlada pelo aluno não melhora a aprendizagem deles.

Por fim, destacamos que conhecer as concepções de avaliação de professores e gestores é igualmente importante.

## REFERÊNCIAS

ATKINSON, P. *Assessment 5-14: What do pupils and parents think?* (Spotlight n. 87). Edinburgh, UK: The SCRE Centre, University of Glasgow, 2003.

BARRETT, P. Structural equation modeling: adjudging model fit. *Personality and Individual Differences*, v. 42, n. 5, p. 815-824, 2007.

BROWN, G. T. L. *Conceptions of assessment: understanding what assessment means to teachers and students*. New York: New Science, 2008. 198 p.

\_\_\_\_\_. *Secondary school students' conceptions of assessment: a survey of four schools*. Conceptions of Assessment and Feedback Project Report #5. Auckland, NZ: University of Auckland, 2006.

\_\_\_\_\_. *Teachers' conceptions of assessment 2002*. 222 f. Tese (Doutorado em Educação) – Faculty of Education, University of Auckland, Auckland, 2002.

BROWN, G. T. L.; HIRSCHFELD, G. H. F. *Secondary school students' conceptions of assessment*. Conceptions of Assessment and Feedback Project Report #4. Auckland: University of Auckland, 2005.

\_\_\_\_\_. Students' conceptions of assessment and mathematics achievement: evidence for the power of self-regulation. *Australian Journal of Educational and Developmental Psychology*, n. 7, p. 63-74, 2007.

\_\_\_\_\_. Students' conceptions of assessment: links to outcomes. *Assessment in Education: principles, policy and practice*, v. 15, n. 1, p. 3-17, 2008.

- BROWN, G. T. L.; IRVING, S. E.; PETERSON, E. R.; HIRSCHFELD, G. H. F. Use of interactive-informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning & Instruction*, v. 19, n. 2, p. 97-111, 2009a.
- BROWN, G. T. L.; PETERSON, E. R.; IRVING, S. E. Self-regulatory beliefs about assessment predict mathematics achievement. In: MCINERNEY, D. M.; BROWN, G. T. L.; LIEM, G. A. D. (Ed.) *Student perspectives on assessment: What students can tell us about improving school outcomes?* Greenwich, CT: Information Age Press, 2009b.
- CAMARGO, A. L. C. O Discurso sobre a avaliação escolar do ponto de vista do aluno. *Revista da Faculdade de Educação*, v. 23, n. 1-2, p. 97-109, jan./dez., 1997.
- CHEUNG, G. W.; RENSVDOLD, R. B. Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, v. 9, n. 2, p. 233-255, 2002.
- CRESWELL, J. W. *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, New Jersey: Pearson Merrill Prentice Hall, 2005. 623 p.
- DILLMAN, D. A. *Mail and internet surveys: the tailored design method*. New York: John Wiley & Sons, 2000. 464 p.
- ENTWISTLE, N. J.; ENTWISTLE, A. Contrasting forms of understanding for degree examinations: the student experience and its implications. *Higher Education*, n. 22, p. 205-227, 1991.
- FAN, X.; SIVO, S. A. Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, v. 42, n. 3, p. 509-529, 2007.
- FLORA, D. B.; CURRAN, P. J. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological methods*, v. 9, n. 4, p. 466-491, 2004.
- HU, L.; BENTLER, P. M. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, v. 6, n. 1, p. 1-55, 1999.
- KLEM, L. Structural equation modeling. In: GRIMM, L. G.; YARNOLD, P. R. (Ed.), *Reading and understanding more multivariate statistics*. Washington, DC: APA, 2000. p. 227-260.
- MARSH, H. W.; HAU, K.; GRAYSON, D. Goodness of fit in structural equation models. In: MAYDEU-OLIVARES, A.; MCARDLE, J. J. (Ed.), *Contemporary psychometrics: a Festschrift for Roderick P. McDonald*. Mahwah, NJ: Lawrence Erlbaum, 2005. p. 275-340.
- MARSH, H. W.; HAU, K.; WEN, Z. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing hu and bentler's (1999) Findings. *Structural Equation Modeling*, v. 11, n. 3, p. 320-341, 2004.
- MATOS, D. A. S. *A avaliação no ensino superior: concepções múltiplas de estudantes brasileiros*, 2010. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

MONI, K. B.; VAN KRAAYENOORD, C. E.; BAKER, C. D. Students' perceptions of literacy assessment. *Assessment in Education*, v. 9, n. 3, p. 319-342, 2002.

PASQUALI, L. (Org.) *Instrumentos psicológicos: manual prático de elaboração*. Brasília: LabPAM, 1999.

REAY, D.; WILLIAM, D. 'I'll be a nothing': structure, agency and the construction of identity through assessment. *British Educational Research Journal*, v. 25, n. 3, p. 343-354, 1999.

STRUYVEN, K.; DOCHY, F.; JANSSENS, S. Students' perceptions about evaluation and assessment in higher education: a review. *Assessment & evaluation in higher education*, v. 30, n. 4, p. 325-341, 2005.

THOMPSON, B. Ten commandments of structural equation modeling. In: GRIMM, L. G.; YARNOLD, P. R. (Ed.), *Reading and understanding more multivariate statistics*. Washington, DC: APA, 2000. cap. 8, p. 261-283.

ZIMMERMAN, B. J. Self-regulated learning and academic achievement: an overview. *Educational psychologist*, v. 25, n. 1, p. 3-17, 1990.

---

#### **DANIEL ABUD SEABRA MATOS**

Doutor em Educação. Professor do Departamento de Educação da Universidade Federal de Ouro Preto (UFOP)  
*danielmatos@ichs.ufop.br; dseabram@yahoo.com.br*

#### **SÉRGIO DIAS CIRINO**

Doutor em Psicologia. Professor da Faculdade de Educação da Universidade Federal de Minas Gerais (UFMG)  
*sergiocirino99@yahoo.com*

#### **GAVIN THOMAS LUMSDEN BROWN**

Doutor em Educação. Professor da Faculty of Education, The University of Auckland  
*gt.brown@auckland.ac.nz*

#### **WALTER LANA LEITE**

Doutor em Educação. Professor do College of Education, University of Florida (UF)  
*walter.leite@coe.ufl.edu*

Recebido em: DEZEMBRO 2011  
Aprovado em: JULHO 2012

## **ANEXO A**

**Itens do questionário *Students' Conceptions of Assessment (SCoA)* version VI, desenvolvido por Brown (2006) e adaptado por Matos (2010)**

1. Eu presto atenção nos meus resultados de avaliação para me concentrar no que eu posso melhorar da próxima vez.
2. A avaliação encoraja a minha turma a trabalhar junta e a ajudar uns aos outros.
3. A avaliação é injusta com os alunos.
4. Os resultados da avaliação mostram o quanto eu sou inteligente.
5. A avaliação ajuda os professores a acompanhar o meu progresso.
6. A avaliação é uma experiência envolvente e agradável para mim.
7. Eu ignoro as informações da avaliação.
8. A avaliação é uma forma de determinar o quanto eu aprendi do ensino.
9. A avaliação averigua o meu progresso em comparação com os objetivos de aprendizagem.
10. Eu faço uso do *feedback* que recebo para melhorar meu aprendizado.
11. A avaliação fornece informação sobre como as escolas estão indo.
12. A avaliação me motiva e aos meus colegas a ajudarem uns aos outros.
13. A avaliação interfere no meu aprendizado.
14. Eu observo o que eu fiz de errado ou de maneira insuficiente para guiar o que eu deveria aprender em seguida.
15. Eu uso as avaliações para assumir responsabilidade para as minhas próximas etapas de aprendizagem.
16. Os resultados da avaliação predizem o meu desempenho futuro.

17. A nossa turma se dá mais apoio quando nós somos avaliados.
18. Os professores avaliam exageradamente.
19. Eu uso as avaliações para identificar o que eu preciso estudar em seguida.
20. A avaliação é importante para a minha carreira futura ou emprego.
21. Quando nós fazemos avaliações, existe um bom clima na nossa turma.
22. Os resultados da avaliação não são muito exatos.
23. Os meus professores usam a avaliação para me ajudar a melhorar.
24. A avaliação mede a qualidade das escolas.
25. A avaliação faz a nossa turma cooperar mais uns com os outros.
26. A avaliação é sem valor.
27. Os professores usam os meus resultados da avaliação para ver o que eles precisam me ensinar em seguida.
28. Quando nós somos avaliados, a nossa turma se torna mais motivada para aprender.
29. Eu desconsidero os meus resultados de avaliação.
30. A avaliação mostra se eu posso analisar e pensar criticamente sobre um assunto.
31. Eu realmente aprecio o aprendizado quando eu sou avaliado.
32. A avaliação tem um impacto pequeno no meu aprendizado.
33. A avaliação diz para os meus pais o quanto eu aprendi.