

Daniel Abud Seabra Matos

**A avaliação no ensino superior: concepções
múltiplas de estudantes brasileiros**

Tese apresentada ao Curso de Doutorado da
Faculdade de Educação da Universidade Federal de
Minas Gerais, como requisito parcial à obtenção do
título de Doutor em Educação

Linha de Pesquisa: Psicologia, Psicanálise e
Educação

Orientador: Dr. Sérgio Dias Cirino

Co-orientador: Dr. Gavin Thomas Lumsden Brown

Belo Horizonte
Faculdade de Educação da UFMG

2010

Universidade Federal de Minas Gerais
Faculdade de Educação
Programa de Pós-graduação em Educação - Conhecimento e Inclusão Social

Tese intitulada “A avaliação no ensino superior: concepções múltiplas de estudantes brasileiros”, de autoria do doutorando Daniel Abud Seabra Matos, aprovada pela banca examinadora constituída pelos seguintes professores:

Prof. Dr. Sérgio Dias Cirino - FAE/UFMG - Orientador

Prof. Dr. José Francisco Soares - FAE/UFMG

Profa. Dra. Elizabeth do Nascimento - FAFICH/UFMG

Prof. Dr. Ricardo Primi - USF

Prof. Dr. Gleiber Couto Santos - UFG

Prof. Dr. Sérgio Luiz Talim - UFMG

Belo Horizonte, 8 de Novembro de 2010.

AGRADECIMENTOS

Ao meu pai, Rui Seabra Matos (*in memoriam*), por ter sido um pai carinhoso, atencioso e um exemplo de vida para mim. Por todo o incentivo que me deu na minha carreira acadêmica e profissional. Graças a você, pai, consegui chegar até aqui por ter me propiciado uma educação de qualidade. Obrigado por tudo.

À minha esposa, Glaucia, pelo apoio incondicional, paciência e carinho nos momentos mais difíceis. Por sua presença sempre confortante.

À minha filha Joana, que nasceu durante o período do Doutorado e trouxe muita felicidade para minha vida. Seu sorriso por muitas vezes me deu força para continuar com esse trabalho.

À Ione (*in memoriam*), Cristina, Luis, Tereza, Maria Borges e toda a minha família.

Ao meu orientador, Sérgio Dias Cirino, por ter me acompanhado durante tantos anos da minha trajetória acadêmica. Por todas suas contribuições nesse período, inclusive o incentivo para o Doutorado Sanduíche.

Special thanks to my supervisor, Professor Gavin T. L. Brown. I would like to thank you for your assistance and feedback through my Ph.D. research. Without you, the three letters after my name would not be possible. Finally, thanks for trusting in me and in my work, and for the opportunities you gave me. I hope to visit you soon. Wait for me. The bus is coming, and the story of how Brazilian students conceive of assessment can be told before it leaves!

Ao amigo Walter Lana Leite, por ter me recebido tão bem no meu período do Sanduíche nos Estados Unidos. Pela disponibilidade para tantas discussões, pela ajuda inestimável com a parte estatística e pela amizade. Sem você essa pesquisa não teria sido possível.

Ao amigo Frederico Assis Cardoso, pela amizade, pelos debates acadêmicos, pelos aprendizados conjuntos, pelos momentos difíceis que compartilhamos, por ter contribuído para que o período do Sanduíche do Doutorado fosse possível.

À professora Helivane de Azevedo Evangelista, pela ajuda inestimável em todas as etapas da pesquisa. Pelas parcerias e pela amizade.

À Anelise Fonseca Dutra e Cristiane Kaitel pela ajuda com esta pesquisa.

Aos professores José Francisco Soares e Regina Helena de Freitas Campos, pelas valiosas sugestões na banca de qualificação do Doutorado.

Ao professor Sérgio Talim, pelas sugestões no projeto do Doutorado.

À Luciana Nocetti Croitor, pela importante ajuda na coleta de dados da pesquisa.

Ao Alexandre Kaitel, pela ajuda com este trabalho, pelos diálogos acadêmicos e profissionais e pela amizade.

À Mary Brito d'Andréa, pela leitura atenta e revisão deste trabalho.

Às professoras do Departamento de Psicologia da UFMG, Larissa Rodrigues e Elizabeth do Nascimento.

À Laura Cançado Ribeiro, pelo apoio e presença em todos os momentos.

À Faculdade de Educação da UFMG, que me acolheu tão bem durante o Mestrado e Doutorado. Aos professores da FAE e aos colegas da pós-graduação.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa do Doutorado Sanduíche que possibilitou o meu período de estudos nos Estados Unidos.

A todos os diretores, coordenadores de curso, professores e alunos que tornaram essa pesquisa possível.

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.

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) - *version VI* (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. Utilizamos como referencial teórico a teoria da aprendizagem auto-regulada. Quatro estudos foram realizados 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). Estudantes de 18 cursos de graduação ($N= 756$; 216 homens e 540 mulheres; idade $M = 24.53$ anos; $DP= 5.91$) de duas IES de Minas Gerais participaram no ano de 2008 (amostragem não-probabilística): uma Universidade Federal ($N= 297$) e um Centro Universitário privado ($N= 459$). As respostas do SCoA brasileiro se ajustaram bem à estrutura fatorial original de oito fatores do questionário ($TLI = .95$; $\gamma \hat{=} .92$; $SRMR = .057$). O presente estudo indicou que duas dimensões (formalidade e locus de controle) e a 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. 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. Tomando a amostra como um todo, os alunos parecem conceber que a avaliação primeiramente melhora o ensino e a aprendizagem e em segundo lugar torna os alunos e escolas responsáveis. A partir da análise das questões abertas do SCoA, elaboramos as seguintes categorias analíticas: a) O papel da avaliação: percepções dos alunos sobre as funções da avaliação; b) Percepções sobre os efeitos da avaliação; c) Percepções sobre a Educação: concepções implícitas nos discursos analisados. Quanto às relações entre as concepções e as definições de avaliação dos estudantes, a definição de avaliação como práticas formais (controle do aluno e do professor) diminuiu com

uma concordância de que a avaliação era uma experiência agradável e algo 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/controle do professor” ainda aumenta com uma concordância de que a avaliação era para a melhora do aluno. Assim, poderíamos acrescentar a seguinte parte na fala hipotética dos alunos: *“Além disso, se a atividade avaliativa é formal e controlada pelo professor, ela me ajuda a melhorar o 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. Dessa forma, os alunos parecem perceber que todos os tipos de atividades avaliativas estão associados a 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. 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”*. Quando comparamos as duas IES, os alunos do centro universitário privado acreditam mais que a avaliação melhora o ensino-aprendizagem e torna alunos e escolas responsáveis do que os alunos da universidade pública. Além disso, os alunos da instituição pública acreditam em um maior grau que a avaliação é algo ruim ou injusto. Por fim, os alunos do centro universitário privado 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.

Palavras-chave: concepções; avaliação; pesquisa de *survey*; análise fatorial confirmatória; alunos; ensino superior; Brasil.

Matos, D. A. S. (2010). *Assessment in Higher Education: multiple conceptions of Brazilian Students*. Doctoral dissertation, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil.

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) - version VI - inventory (Brown, 2006) for the Brazilian context; (3) establish the underlying dimensions shaping Brazilian university students' definitions of assessment; (4) analyze how conceptions of assessment predict students' definitions of assessment; and (5) investigate differences between students at public and private universities. We used as theoretical framework the theory of self-regulated learning. Four studies were carried out using a mixture of quantitative (multidimensional scaling, exploratory and confirmatory factor analysis, analysis of variance, structural equation modeling) and qualitative (content analysis was used to investigate the open-ended questions in the survey) methods. Students enrolled in 18 undergraduate courses ($N = 756$, 216 males and 540 females, age $M = 24.53$ years, $SD = 5.91$) from two universities in the state of Minas Gerais participated in 2008 year (nonrandom sample): a public university ($N = 297$) and a private University ($N = 459$). The Brazilian SCoA responses fit well to the original eight factors structure of the questionnaire ($TLI = .95$; $\gamma^2 = .92$, $SRMR = .057$). Students defined assessment in terms of their formality and who was in control of them, resulting in four types of assessment practices: informal student-controlled; informal teacher-controlled; formal student-controlled; and formal teacher-controlled. The students emphasized conventional, formal evaluation methods controlled by the teacher. Even supposedly student-controlled practices (i.e., self and peer-assessments) were perceived as teacher-controlled practices. As a whole, the university students conceived assessment most strongly as being about improvement, and second as about accountability. From the analysis of the open-ended questions attached to the SCoA, three analytical categories arose: a) perceptions about the functions or roles of assessment; b) perceptions about assessment effects; and c) perceptions about education itself. The definitions of assessment were meaningfully predicted by student responses to key factors in the SCoA inventory. The SCoA factor of "Improvement" was only associated with formal practices and those controlled by the teacher. Formal practices, whether controlled by students or teachers, decreased as agreement with assessment as an enjoyable experience and something bad. In other words, we might hear students say something like: *"If assessment is formal, I do not necessarily think it's an enjoyable*

experience for me. Nevertheless, I also do not consider formal assessment is bad". The formal teacher-controlled definition also increased with agreement that assessment was for improving students' learning, suggesting that Brazilian university students additionally think: *"If the assessment practice is formal and controlled by the teacher, it helps me improve my learning; it helps me in my next steps."* In contrast, all definitions of assessment, both formal and informal, were positively predicted by endorsement of the concept that assessment leads to benefits in motivation of classmates and climate of the classroom. Finally, the definition of assessment as informal practices controlled by the teacher decreased with increased endorsement of assessment as something to be ignored, suggesting students think: *"If assessment is informal and controlled by the teacher, it's good for class dynamics and climate of the classroom, and I do not ignore this type of assessment activity."* When comparing the two universities, the private university students endorsed more that assessment improves teaching-learning and makes students and schools accountable than the public university students. Furthermore, the public university students agreed more that assessment was bad or unfair, while the private university students perceived assessment as a more enjoyable experience. Finally, the private university students associated more strongly assessment as improvement and affective benefits such as the motivation of classmates and climate of the classroom.

Key words: conceptions; assessment; survey research; confirmatory factor analysis; students; higher education; Brazil.

LISTA DE ILUSTRAÇÕES

Figura 1 - O Sistema Nacional de Avaliação da Educação Superior (SINAES).....	27
Figura 2 - Níveis de mensuração das variáveis.....	70
Figura 3 - Esquema dos estudos desenvolvidos na pesquisa.....	77
Figura 4 - Mapa perceptual: tipos de atividades avaliativas.....	83
Figura 5 - Médias das respostas das concepções de avaliação.....	98
Figura 6 - Modelagem de Equações Estruturais.....	108
 Quadro 1 - Definições de concepção.....	 35

LISTA DE TABELAS

1 - Matrículas em Cursos de Graduação Presenciais: sexo dos matriculados em 2008.....	76
2 - Tipos de atividades avaliativas e os quadrantes do Escalonamento Multidimensional.....	84
3 - Comparação de índices de ajuste dos modelos da AFC: Definições de avaliação.....	88
4 - Cargas fatoriais padronizadas dos itens do SCoA Brasileiro: Definições de avaliação.....	89
5 - Comparação da consistência interna do SCoA entre amostras do Brasil e Nova Zelândia.....	90
6 - Correlação entre os fatores do SCoA Brasileiro.....	91
7 - Comparação de índices de ajuste dos modelos da AFC: Concepções de avaliação.....	92
8 - Cargas fatoriais padronizadas dos itens do SCoA Brasileiro: Concepções de avaliação.....	93
9 - Comparação de índices de ajuste da AFE: Concepções de avaliação.....	95
10 - AFE: cargas fatoriais dos itens do SCoA Brasileiro - Concepções de avaliação.....	96
11 - Diferenças entre IES pública e IES privada nos fatores do SCoA.....	99
12 - Teste robusto de igualdade de médias: tipo de IES.....	100

SUMÁRIO

1 ELEMENTOS PARA UMA ANÁLISE DA AVALIAÇÃO NO CONTEXTO DO ENSINO SUPERIOR.....	14
1.1 A avaliação sob o ponto de vista dos estudantes.....	14
1.2 Reformas na área da avaliação e sua relação com o ensino público e privado	16
1.3 Legislação e avaliação educacional.....	22
1.4 O Sistema Nacional de Avaliação da Educação Superior (SINAES).....	26
2 REVISÃO DA LITERATURA.....	32
2.1 Concepções de avaliação.....	32
2.1.1 A questão da terminologia nos estudos sobre as concepções de avaliação.....	34
2.2 As concepções de avaliação dos alunos universitários brasileiros: uma revisão de estudos.....	36
2.2.1 Estudos com poucas concepções.....	38
2.2.2 Estudos com concepções múltiplas.....	42
2.2.3 Tipos de atividades avaliativas.....	43
2.2.4 Uma análise sobre a revisão de estudos brasileiros.....	44
3 CARACTERIZAÇÃO E ADAPTAÇÃO DO <i>STUDENTS' CONCEPTIONS OF ASSESSMENT</i> (SCOA).....	49
3.1 Uma descrição do SCoA.....	49
3.2 Referencial teórico do SCoA.....	52
3.2.1 Aprendizagem auto-regulada.....	52
3.3 O contexto e resultados de pesquisas desenvolvidas na Nova Zelândia.....	56
3.4 Diferenças culturais, tradução e validação de instrumentos.....	58
4 OBJETIVOS E HIPÓTESES DA PESQUISA	61

4.1 Objetivos.....	61
4.1.1 Objetivo Geral.....	61
4.1.2 Objetivos específicos.....	61
4.2 Hipóteses.....	61
5 MÉTODO	63
5.1 O desenho de pesquisa.....	63
5.1.1 Escalas de resposta.....	67
5.2 Métodos estatísticos.....	69
5.2.1 Níveis de mensuração das variáveis.....	69
5.2.2 Escalonamento Multidimensional.....	71
5.2.3 Análise Fatorial.....	72
5.2.4 Modelagem de Equações Estruturais.....	73
5.2.5 Qualidade de ajuste: índices de ajuste do modelo.....	74
5.3 Análise de conteúdo.....	75
5.4 Caracterização da amostra.....	75
5.5 Etapas da pesquisa	77
5.5.1 Procedimentos teóricos.....	78
5.5.2 Procedimentos empíricos (experimentais)	79
6 RESULTADOS E DISCUSSÃO	80
6.1 Adaptação e validação do SCoA.....	80
6.2 Definições de avaliação.....	82
6.2.1 Escalonamento Multidimensional.....	82
6.2.2 Análise Fatorial Confirmatória.....	87
6.3 Concepções de avaliação.....	90
6.3.1 Consistência interna e evidências de validade.....	90

6.3.2 Análise Fatorial Confirmatória.....	92
6.3.3 Análise Fatorial Exploratória.....	94
6.3.4 Análises adicionais sobre as concepções de avaliação.....	97
6.3.5 Diferenças entre a IES pública e a IES privada.....	98
6.4 Resultados qualitativos.....	101
6.4.1 Análise de conteúdo.....	103
6.5 Relação entre as concepções e as definições de avaliação: Modelagem de Equações Estruturais.....	107
6.6 Considerações finais.....	111
 REFERÊNCIAS	 113
 ANEXO A - Questionário <i>Students' Conceptions of Assessment</i> (SCoA - VI).....	 124
ANEXO B - Quadro completo do processo de tradução do SCoA.....	128
ANEXO C - Roteiro de entrevista semi-estruturada sobre o SCoA.....	136
ANEXO D - Declaração de autorização de pesquisa.....	137
ANEXO E - Termo de consentimento livre e esclarecido.....	138

CAPÍTULO 1 - ELEMENTOS PARA UMA ANÁLISE DA AVALIAÇÃO NO CONTEXTO DO ENSINO SUPERIOR

1.1 - A avaliação sob o ponto de vista dos estudantes

Na década de 1990, um cenário nacional e internacional destacou a centralidade da avaliação educacional e desde essa época está acontecendo uma retomada da avaliação no Brasil. Esse aumento do interesse na avaliação educacional também produziu um aumento do número de trabalhos sobre a avaliação. Na literatura especializada brasileira, são frequentes estudos sobre assuntos como: avaliação de políticas públicas educacionais (SOUZA; VASCONCELOS, 2006; WERLE, 2006); avaliação de projetos sociais e educacionais (DÓRIA; TUBINO, 2006); avaliação do ambiente de aprendizagem da sala de aula (MATOS, 2006; MATOS *et al.*, 2008); avaliação de docentes do Ensino Superior (BOCLIN, 2004); avaliação do desempenho cognitivo dos alunos (SOARES *et al.*, 2001); avaliação de inovações educacionais (BORGES; GONÇALVES; CUNHA, 2003); avaliações sobre a qualidade do sistema educacional brasileiro por meio de exames como o Exame Nacional de Desempenho dos Estudantes - ENADE (VERHINE *et al.*, 2006); entre outros.

Dentre as várias possibilidades de abordagem da avaliação, a análise das concepções de avaliação dos alunos tem sido uma das opções de trabalho no campo educacional. No entanto, existem poucas pesquisas que consideram a avaliação sob o ponto de vista dos próprios estudantes. Considerando os diversos níveis de ensino, as pesquisas sobre as concepções de avaliação dos alunos são escassas (BROWN; HIRSCHFELD, 2005). Além disso, a literatura disponível sobre as concepções de avaliação dos alunos é geralmente focada no ensino médio¹ ou no ensino superior.

As concepções de avaliação dos alunos são especialmente importantes porque: a avaliação tem um papel essencial na qualidade do aprendizado; as crenças e atitudes dos estudantes com relação aos propósitos da avaliação parecem influenciar as suas práticas e resultados (ENTWISTLE; ENTWISTLE, 1991). Dessa forma, a maioria das pesquisas descreve que as concepções dos alunos sobre a avaliação influenciam de maneira significativa a abordagem deles com relação ao aprendizado (STRUYVEN; DOCHY; JANSSENS, 2005).

A literatura da área também evidencia associações entre as concepções de avaliação e o desempenho acadêmico (notas) dos alunos (BROWN; HIRSCHFELD, 2007, 2008). Por exemplo: a teoria da aprendizagem auto-regulada (referencial teórico

¹ Nessa pesquisa, optamos pela seguinte tradução quanto aos níveis de ensino: *primary education*: ensino fundamental; *secondary education*: ensino médio; *tertiary/ higher education / college*: ensino superior.

da presente pesquisa) indica que alunos que assumem responsabilidade por suas próprias ações geralmente atingem melhores resultados na maioria das medidas educacionais. Já aqueles que não assumem responsabilidade por suas ações tendem a atingir piores resultados. Estudantes que percebem a avaliação como uma força construtiva para a responsabilidade pessoal obtêm melhores notas. Já alunos que procuram responsabilizar escolas ou professores pelos resultados avaliativos, que não assumem a avaliação seriamente, ou que a ignoram têm notas piores (BROWN; HIRSCHFELD, 2005).

Considerando a avaliação no ensino superior brasileiro, os alunos vêm sendo expostos a um contexto de diversas reformas e mudanças: novas legislações, novas práticas avaliativas, novos formatos de cursos e de Instituições de Ensino Superior (IES), novas exigências do mercado, entre outras. A chamada avaliação de sala de aula, que sempre fez parte da vida dos estudantes, vem passando por modificações: maior uso de avaliação formativa; implementação gradativa de novas práticas avaliativas (exemplos: autoavaliação, portfólio); maior ênfase no caráter transdisciplinar do conhecimento científico. A avaliação institucional também tem sido cada vez mais incorporada à realidade das IES: os alunos têm o ENADE como componente curricular obrigatório (devem demonstrar conhecimento em uma avaliação externa) e precisam acompanhar a avaliação da IES na qual estudam para saber o conceito dela. A avaliação da sua instituição (via credenciamento e credenciamento) ou do seu curso (via autorização, reconhecimento e renovação de reconhecimento) também pode afetar a sua vida. Esses são apenas alguns exemplos das complexas transformações que estão acontecendo. Todos esses processos de transformação da avaliação educacional são colocados em prática e vivenciados por diversos atores sociais: corpo discente e docente das IES, funcionários técnico-administrativos, reitores, sindicatos, parlamentares, representantes do governo e sociedade civil. Contudo, o foco desse trabalho será o ponto de vista dos estudantes.

Nesse sentido de mudanças, tradicionalmente os alunos universitários eram avaliados de maneira somativa com provas no final das disciplinas. Com as transformações que vêm ocorrendo (principalmente o maior uso de avaliação formativa), os estudantes têm sido participantes mais ativos nas avaliações que acontecem em sala de aula. Isso envolve práticas avaliativas não tradicionais como a avaliação feita pelos próprios colegas (pares) e a autoavaliação. Entretanto, a eficácia dessas práticas avaliativas tidas como inovadoras está condicionada pela 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 importante. Assim, destacamos uma questão central para a presente pesquisa: como os estudantes universitários estão percebendo a avaliação no contexto atual ?

Apesar de a nossa pesquisa estar centrada principalmente nas concepções dos alunos sobre práticas avaliativas de sala de aula, consideramos que tais práticas são consequência de um contexto mais amplo de reformas e mudanças na avaliação. Dessa forma, apresentamos nesse capítulo alguns elementos para uma análise da avaliação no ensino superior brasileiro em um contexto mais amplo, a saber: reformas na área da avaliação e sua relação com o ensino público e privado (um dos nossos objetivos específicos é comparar as concepções de avaliação de alunos pertencentes a uma IES pública e a uma IES privada); mudanças na legislação da avaliação educacional; atual forma de avaliação das IES brasileiras (Sistema Nacional de Avaliação da Educação Superior - SINAES).

1.2 - Reformas na área da avaliação e sua relação com o ensino público e privado

O aumento significativo do número de trabalhos teóricos sobre a avaliação reflete uma tendência mundial, pois as mudanças no ensino superior brasileiro acompanham tendências internacionais. Estudos de caráter comparativo demonstram que países de realidades diferentes têm percorrido caminhos muito semelhantes no desenvolvimento das políticas de avaliação, que se caracterizam por: crescente controle nacional sobre os processos avaliativos; certa uniformidade de objetivos e práticas; maior frequência de ações avaliativas e aumento do seu campo de intervenção (AFONSO, 2002).

Devido à globalização e ao capital mundializado, o ensino superior é submetido a essa regulação de caráter internacional. Os mecanismos de controle têm aumentado em quase todos os países, sendo a comparação entre nações uma exigência desse processo. A educação passa a se submeter às leis de mercado e aos organismos internacionais de regulação do comércio (ex: Organização Mundial do Comercio - OMC) (SOBRINHO, 2006). Essas semelhanças em alguns princípios orientadores das políticas educacionais, no entanto, não significam negar que existem especificidades nacionais (AFONSO, 2002).

Sobrinho (2006) aponta que, para entender a avaliação do ensino superior em processo nos últimos anos no Brasil, precisamos compreender algumas relações estabelecidas entre a avaliação, a educação superior e o Estado. Como resultado dessas relações, dois conceitos que parecem opostos combinam-se nas Instituições de Ensino Superior (IES): autonomia e controle. Ou seja: ao mesmo tempo em que as IES

possuem hoje uma maior autonomia quanto aos processos e meios (ex: modelos de gestão semelhante ao de empresas, facilidade para o surgimento de novos cursos e expansão da iniciativa privada), o contraponto é o controle, a avaliação. Nesse sentido, o autor indica que avaliação e privatização são um eixo central das reformas colocadas em prática por órgãos governamentais e agências multilaterais, destacando o caráter político da avaliação e sua vinculação com o poder.

O tema privatização nos remete a uma discussão mais aprofundada sobre a avaliação nas IES públicas e privadas. Nessa direção, boa parte da literatura brasileira tem sido bastante crítica com relação à avaliação, enfatizando principalmente sua dimensão política.

Calderón (2000), por exemplo, aponta que o ensino superior brasileiro foi alterado de forma significativa a partir do início da década de 1990² pelo aparecimento de instituições denominadas “universidades mercantis”. Essas IES se caracterizam por serem privadas, com explícitos fins lucrativos, geridas como empresas educacionais, oferecendo produtos e serviços de acordo com a demanda do mercado. Assim, os alunos passam a ser tratados como “clientes-consumidores” no campo acadêmico-universitário. Nesse contexto de mercantilização do ensino, Silva (2008) demonstra que o ensino superior privado no Brasil é hoje um objeto de desejo de investidores, pois movimenta mais de US\$ 8.000.000.000 (oito bilhões de dólares) por ano. Para o autor, mesmo com a crise provocada pela abertura indiscriminada de IES privadas (mais oferta de vagas do que demanda), muitas instituições conseguiram se consolidar e passaram a uma fase de expansão seletiva, inclusive incorporando alguns concorrentes. Apesar da grande inadimplência de alunos e do alto número de vagas ociosas, tem acontecido um crescente envolvimento de bancos e outras empresas no ensino superior privado nos últimos anos.

Para Sguissardi (2005), a partir de 1994 se inicia no Brasil um processo de redução de gastos públicos com Instituições Federais de Ensino Superior (IFES). O autor também destaca a retomada (de uma forma muito mais forte) de um processo semelhante ao que havia acontecido na década de 1970 durante a ditadura militar: a privatização do ensino superior. Vieira (2008) sintetiza da seguinte maneira as reformas ocorridas na década de 1990 no ensino superior brasileiro: grande expansão do setor privado; diminuição dos investimentos públicos; aumento significativo de matrículas de

² Analisamos períodos históricos e legislações educacionais relevantes para a contextualização da presente pesquisa. Para uma discussão detalhada sobre esses aspectos, ver Bonamino (2002), Vieira (2008) e Guimarães (2003).

alunos, principalmente no setor privado; aumento do controle do Estado sobre o produto da educação (avaliação).

Esse maior controle do Estado sobre os produtos educacionais geralmente é interpretado na literatura por diversos autores dentro de uma ótica neoliberal: isso representa uma menor presença do Estado na economia e mais controle do campo social. Assim, ocorre um recuo do Estado no que se refere ao plano econômico, enquanto a avaliação cumpriria um papel de jogar a educação no terreno da mercadorização (SOBRINHO, 2006).

Essa questão do financiamento possui uma relação muito próxima com a privatização e a avaliação. Segundo Sobrinho (2006), as IES privadas encontram ao mesmo tempo facilidade para expansão e uma dura competição. Já o setor público vem sofrendo cortes no seu orçamento e se vê obrigado a procurar por fundos privados, como a cobrança por serviços e a parceria com empresas. No contexto de crise de financiamento e expansão da rede privada do ensino superior, Estado e mercado necessitam diversificar as instituições e suas ofertas para absorver demandas diversas. Considerando uma exigência internacional de maior escolaridade e capacitação para o mundo do trabalho, surgem organizações novas e competitivas e instituições antigas se transformam. Muitas vezes uma graduação tradicional não é suficiente para garantir competitividade em um determinado setor, por isso surgem novas modalidades, como cursos de curta duração com finalidades utilitárias ou até mesmo a criação de IES centradas em um único ramo de formação profissional. Outros exemplos nas IES públicas seriam a criação cursos de Pós-Graduação *Lato Sensu* pagos, de fundações, dentre outros. Dessa forma, o autor conclui que a privatização da educação superior se torna cada vez mais forte e presente, levando mesmo o sistema público de ensino a incorporar os mecanismos de planejamento e de medição utilizados no setor privado.

Nesse contexto, o chamado “Estado Avaliador” tem ocupado um papel de destaque. A expressão “Estado Avaliador” é usada hoje para definir um Estado que não se responsabiliza pelos processos, mas que tenta controlar o sistema pelos resultados (por meio da avaliação) (VIEIRA, 2008). Essa expressão significa, em um sentido amplo, que o Estado passa a admitir a lógica do mercado, por meio da importação para o domínio público de modelos de gestão privada, enfatizando os resultados ou produtos dos sistemas educacionais. A avaliação aparece então como um pré-requisito para a implantação de mecanismos de controle sobre os produtos educacionais (AFONSO, 2002). Portanto, a marca registrada das políticas avaliativas do Estado Avaliador tem sido a fiscalização, o controle e a responsabilização (VIEIRA, 2008).

Desde a década de 1980, aconteceu um aumento da chamada “responsabilização” (*accountability*)³, que se tornou um mecanismo avaliativo de destaque. Como já colocamos anteriormente, dentro da lógica do Estado Avaliador, o controle e a fiscalização recaem principalmente sobre os resultados finais. Os governos têm exigido das IES a demonstração de forma objetiva dos seus graus de eficiência e produtividade (em outras palavras, devem prestar contas). Dessa forma, a qualidade se torna quase um sinônimo de eficiência sendo apreendida principalmente de uma maneira quantitativa. Assim, indicadores quantificáveis da produtividade das IES passam a ser relacionados com a responsabilização. Uma das consequências disso é o recebimento de benefícios ou recursos por parte da instituição (SOBRINHO, 2006). Outro exemplo de consequência pode ser o descredenciamento da IES.

Ainda com relação à avaliação no ensino superior, Chaui (2003) indica uma série de pontos para reflexão. São alguns desses elementos: a) a universidade, reduzida a uma organização, abandona a formação e a pesquisa para entrar em uma fragmentação competitiva. A avaliação do seu trabalho é feita em termos de custo benefício, sendo pautada pela ideia de produtividade (avaliação do tempo, do custo e da quantidade de produção). Isso acontece porque a maior parte de suas pesquisas sofre determinações do mercado e dos financiadores. Em outras palavras, a universidade pública produz um conhecimento destinado à apropriação privada, como consequência da mudança significativa sofrida pelas ciências em sua relação com a prática: a ciência tornou-se uma força produtiva e se transformou num componente do próprio capital; b) os organismos internacionais que discutem o ensino superior têm enfatizado a chamada “sociedade do conhecimento”, que implica em velocidade, em uma grande redução do tempo entre a aquisição de um conhecimento e sua aplicação tecnológica. Essa “explosão do conhecimento” gera uma série de consequências com relação à avaliação. Os processos de avaliação da produção acadêmica, que influenciam a carreira do docente e a obtenção de financiamento de pesquisas, são baseados na quantidade de

³ O termo *accountability* tem sido traduzido na literatura como responsabilização, mas é de difícil tradução porque em inglês essa palavra carrega um sentido repleto de informações que precisam ser explicitadas em português. *Accountable* pode ser traduzido como sujeito à obrigação de reportar, explicar ou justificar algo a um terceiro. Segundo Schedler (1999), *accountability* é definido como: A é *accountable* para B quando A é obrigado a informar a B sobre suas decisões e ações, para justificá-las, sendo que A pode sofrer punições no caso de condutas inadequadas. Na educação, pode ser definida como uma política de tornar escolas e professores responsáveis pelo progresso acadêmico dos alunos ligando esse progresso com recursos para salários, manutenção das escolas, entre outros. Um exemplo de *accountability* é a avaliação institucional que acontece no ensino superior do Brasil como um dos componentes do Sistema Nacional de Avaliação da Educação Superior (SINAES). De acordo com os resultados nesse sistema de avaliação, a Instituição de Ensino Superior (IES) pode receber, por exemplo, a autorização para funcionar como Faculdade, Centro Universitário ou Universidade. Isso tem como consequência um grau de autonomia menor ou maior no funcionamento da IES.

publicação de artigos e de participações em congressos. A quantidade de pontos obtidos pelo pesquisador depende de onde ele consegue publicar seus artigos, pois os periódicos científicos são definidos hierarquicamente por um *ranking*. A avaliação passou a ser pautada por critérios de eficácia e de competitividade, deixando cada vez mais de ser feita pelos pares. Chaui destaca que a quantidade de publicações precisa ser vista com cautela, pois essa quantidade pode significar pouca qualidade e pouca inovação. Todos esses elementos remetem a uma questão de fundo: a mudança imposta ao tempo do trabalho intelectual e científico (diminuição desse tempo).

Outro fenômeno discutido na literatura sobre a avaliação no ensino superior é a chamada desprofissionalização do magistério superior. Nesse fenômeno, existe uma complementaridade entre avaliação e uma mentalidade empresarial. Assim, as IES atuam como empresas reduzindo o número de professores em regime de dedicação exclusiva, precarizando os contratos de trabalho dos professores, prestigiando docentes que respondem a demandas mercadológicas, dentre outros. A avaliação cumpriria um papel de consolidar essa mentalidade verificando o desempenho e contabilizando os produtos. Existe hoje uma tendência em diversos países que pode ser caracterizada como a desprofissionalização do magistério superior: as IES passam cada vez mais a cumprir uma função reduzida e utilitária de capacitação profissional de acordo com as habilidades e competências exigidas pelo mercado de trabalho. Com isso, a docência do ensino superior entra em uma crise (SOBRINHO, 2006).

Portanto, é impossível desvincular a discussão sobre a avaliação no ensino superior da discussão sobre ensino público e privado. Além disso, a literatura descreve de forma bastante crítica a avaliação no ensino superior do Brasil, destacando tanto um caráter político quanto uma vinculação com o poder. Sampaio (1998) sintetiza essas questões, apontando que boa parte da literatura brasileira evidencia uma visão negativa sobre o ensino superior privado. A autora ainda afirma que, em geral, os estudos compartilham os seguintes pontos de vista: a oferta da iniciativa privada é vista como algo indesejado (o Estado deveria universalizar o ensino superior público e gratuito); na atual conjuntura democrática, a política dos governos militares responsável pela expansão do ensino privado é substituída pela política neoliberal guiada pelos organismos internacionais; o ensino superior privado é guiado pela lógica da mercantilização do ensino e funciona, geralmente, como uma fábrica de diplomas.

Apesar de concordarmos com diversas críticas que são feitas (como a dimensão política da avaliação), acreditamos que a discussão sobre a avaliação tem sido

apresentada de uma forma parcial, centrada apenas nos aspectos negativos. Essa temática complexa por vezes é simplificada.

Em primeiro lugar, mesmo com todos os problemas relacionados à avaliação e a forma como ela tem sido conduzida no ensino superior, a literatura em geral não aponta nenhum tipo de benefício ou melhoria como consequência da avaliação. Segundo Jorba e Sanmartí (2003), a avaliação apresenta basicamente duas funções: uma de caráter social, de seleção e classificação; outra de caráter pedagógico, de ajuste do processo ensino-aprendizagem, de reconhecimento de mudanças que podem ser introduzidas progressivamente num determinado processo. A segunda função tem uma dimensão formativa, fornecendo informações úteis para a adaptação das atividades de ensino-aprendizagem. Assim, raramente as pesquisas apresentam melhorias associadas à avaliação. Uma exceção, por exemplo, é o trabalho de Felix (2008) que pesquisou IES no Brasil e em Portugal para investigar os efeitos e as mudanças provocadas pelos processos de avaliação institucional externa. Como resultado, a autora destaca que foram introduzidas melhorias nos cursos das IES em consequência das recomendações dos processos avaliativos externos. Nesse sentido, mesmo que as práticas atuais de avaliação no ensino superior careçam de um grau formativo maior, nos perguntamos se nenhum tipo de benefício acontece em função da avaliação.

Em segundo lugar, quanto à relação entre avaliação e ensino público e privado, problematizamos a dificuldade em pensar duas categorias gerais e abstratas denominadas “ensino superior público” e “ensino superior privado”. Acreditamos que as IES públicas apresentam muitas diferenças entre si, assim como as IES privadas. Nesse sentido, a generalização geralmente presente na literatura pode ser perigosa. Outra generalização perigosa presente na literatura se refere aos métodos quantitativos, que são apresentados como acríticos, como exigências de uma política neoliberal (para uma discussão detalhada sobre métodos qualitativos e quantitativos ver seção 2.2.4).

Por fim, destacamos que não estamos adotando aqui uma visão acrítica ou ingênua sobre a avaliação. Acreditamos que apresentar a avaliação no ensino superior apenas como uma política neoliberal é uma simplificação de um tema muito complexo. Não discordamos do caráter político e das relações de poder presentes na avaliação. Mas precisamos avançar na discussão da avaliação no ensino superior, inclusive com relação a propostas de mudanças. Assim, consideramos que investigar como os alunos do ensino superior concebem a avaliação pode nos indicar caminhos e enriquecer essa discussão. Lembramos que os estudantes são atores sociais geralmente pouco ouvidos com relação à temática da avaliação.

1.3 - Legislação e avaliação educacional

A Constituição Federal de 1988 foi consequência de um longo processo de redemocratização e representa um marco na história brasileira. Silva (2008) destaca que a constituição, além dedicar um capítulo especificamente para a Educação, também aborda a temática educacional em várias partes, tanto direta quanto indiretamente.

Segundo Guimarães (2003), a Constituição introduz uma proposta de avaliação da qualidade da educação de toda a rede de ensino superior pública e privada. O art. 209 da Constituição Federal de 1988 determina que “O ensino é livre à iniciativa privada, atendidas as seguintes condições: I - cumprimento das normas gerais da educação nacional; II - autorização e avaliação de qualidade pelo Poder Público”. Já o art. 206 dispõe: “O ensino será ministrado com base nos seguintes princípios: [...] III - pluralismo de ideias e de concepções pedagógicas, e coexistência de instituições públicas e privadas de ensino; [...] VII - garantia de padrão de qualidade”. Nesses dois artigos, dois pontos merecem destaque: a determinação de que “o ensino é livre à iniciativa privada”, da “coexistência de instituições públicas e privadas de ensino”, fortalece um modelo expansionista do ensino superior privado; o contraponto dessa maior liberdade é a avaliação: além do cumprimento de normas gerais, as instituições educacionais precisam ser autorizadas e avaliadas pelo poder público. Assim, a avaliação cumpre um papel fundamental nessa “garantia de padrão de qualidade”. Esse tema se repete em outros artigos. O art. 211 diz que:

Art. 211. A União, os Estados, o Distrito Federal e os Municípios organizarão em regime de colaboração seus sistemas de ensino.

§ 1º A União organizará o sistema federal de ensino e o dos Territórios, financiará as instituições de ensino públicas federais e exercerá, em matéria educacional, função redistributiva e supletiva, de forma a garantir equalização de oportunidades educacionais e **padrão mínimo de qualidade do ensino** mediante assistência técnica e financeira aos Estados, ao Distrito Federal e aos Municípios;
(grifo do autor)

O § 3º do art. 212 determina que “A distribuição dos recursos públicos assegurará prioridade ao atendimento das necessidades do ensino obrigatório, no que se refere a universalização, garantia de padrão de qualidade e equidade, nos termos do plano nacional de educação”. Já o art. 214 dispõe sobre o plano nacional de educação:

Art. 214. A lei estabelecerá o plano nacional de educação, de duração decenal, com o objetivo de articular o sistema nacional de educação em regime de colaboração e definir diretrizes, objetivos, metas e estratégias de implementação para assegurar a manutenção e desenvolvimento do ensino em seus diversos níveis, etapas e modalidades por meio de ações integradas dos poderes públicos das diferentes esferas federativas que conduzam a:

I - erradicação do analfabetismo;

II - universalização do atendimento escolar;

III - melhoria da qualidade do ensino;

IV - formação para o trabalho;

V - promoção humanística, científica e tecnológica do País.

VI - estabelecimento de meta de aplicação de recursos públicos em educação como proporção do produto interno bruto.

(grifo do autor)

Esses artigos revelam que o Brasil acompanhava uma tendência internacional no que se refere a esse binômio qualidade do ensino - avaliação. Afonso (2002), por exemplo, apresenta um boletim da Organização para Cooperação e Desenvolvimento Econômico (OCDE) de 1991 que destacou na época uma renovação em diferentes países do interesse pela avaliação por diversas razões: necessidade dos países possuírem mão de obra qualificada; necessidade de melhorar a qualidade da educação (ponto que acabamos de discutir); nova distribuição de responsabilidades na relação entre autoridades centrais e locais na gestão das instituições escolares (é provável que esse ponto encontre um paralelo com as novas relações estabelecidas entre a União, Estados e Municípios na Constituição de 1988 e em outros documentos importantes como a LDB de 1996).

Outra proposta inovadora da Constituição Federal de 1988 diz respeito à autonomia das universidades. O art. 207 determina que “As universidades gozam de autonomia didático-científica, administrativa e de gestão financeira e patrimonial, e obedecerão ao princípio de indissociabilidade entre ensino, pesquisa e extensão”. Essa autonomia conferida às universidades poderia interessar mais diretamente ao setor privado, pois seria um caminho para expansão com menor controle do MEC (desde que as IES privadas alcançassem o status de universidade). A autonomia seria alcançada por IES dispostas a cumprir alguns requisitos como, por exemplo, a viabilização da pesquisa, investimento na formação do corpo docente e adoção de regime de tempo integral. Essas características estão presentes de maneira muito forte nas IES públicas, ao contrário da realidade da maioria das IES privadas (GUIMARÃES, 2003). No entanto, destacamos que atualmente, na prática existe um certo grau de autonomia

mesmo em IES que não são universidades⁴. Por exemplo: uma IES que atinge o status de Centro Universitário já possui autonomia para criar e extinguir cursos em sua sede e realizar reformas curriculares. O art. 28 do Decreto 5.773 de 2006 referenda essa questão ao determinar que Universidades e Centros Universitários “nos limites de sua autonomia, [...] independem de autorização para funcionamento de curso superior, devendo informar à Secretaria competente os cursos abertos para fins de supervisão, avaliação e posterior reconhecimento [...]”. Para Silva (2008), uma das mudanças mais importantes como relação à organização acadêmica das IES foi a criação dos Centros Universitários que, juntamente com as Universidades, gozam de diversas atribuições relativas à autonomia.

Passada a euforia da Constituição de 1988 e diante da necessidade de regulamentação dos avanços alcançados, ficou evidente que a reforma necessitava de esforços adicionais para ser viabilizada. Assim, na década de 1990, novas legislações vieram com propostas de mudanças no ensino superior. Tais legislações estavam baseadas principalmente em dois eixos: avaliação e autonomia (GUIMARÃES, 2003).

A lei 9.131 de 1995 estabeleceu avaliações periódicas das IES e criou o Exame Nacional de Cursos - ENC (o antigo “provão”):

Art. 3º [...] o Ministério da Educação e do Desporto fará realizar avaliações periódicas das instituições e dos cursos de nível superior, fazendo uso de procedimentos e critérios abrangentes dos diversos fatores que determinam a qualidade e a eficiência das atividades de ensino, pesquisa e extensão.

§ 1º Os procedimentos a serem adotados para as avaliações a que se refere o caput incluirão, necessariamente, a realização, a cada ano, de exames nacionais com base nos conteúdos mínimos estabelecidos para cada curso, previamente divulgados e destinados a aferir os conhecimentos e competências adquiridos pelos alunos em fase de conclusão dos cursos de graduação. [...]

Art. 4º Os resultados das avaliações referidas no § 1º do art. 2º serão, também, utilizados pelo Ministério da Educação e do Desporto para orientar suas ações no sentido de estimular e fomentar iniciativas

⁴ O Decreto 3.860 de 2001 estabelece no art. 7: “Quanto à sua organização acadêmica, as instituições de ensino superior do Sistema Federal de Ensino, classificam-se em: I - universidades; II - centros universitários; e III - faculdades integradas, faculdades, institutos ou escolas superiores”. Resumidamente, as diferenças entre Faculdade, Centro Universitário e Universidade são: as Faculdades e institutos superiores não possuem nenhuma autonomia e precisam sempre recorrer ao MEC para a autorização de qualquer atividade. Oferecem cursos de graduação e pós-graduação na modalidade *lato sensu*. Seus diplomas precisam ser registrados por uma universidade. Os Centros Universitários gozam de certa autonomia. Não são obrigados a fazer pesquisa e também não são obrigados a oferecer pós-graduação *stricto sensu*. As universidades possuem autonomia plena e não precisam de autorização do MEC para criar/fechar cursos e sedes, aumentar ou diminuir vagas, expedir diplomas. As Universidades precisam oferecer: formação na graduação; pesquisa e extensão; cursos em pelo menos cinco áreas do conhecimento.

voltadas para a melhoria da qualidade do ensino, principalmente as que visem a elevação da qualificação dos docentes.⁵

A Lei 9.394 (Lei de Diretrizes e Bases da Educação Nacional - LDB de 1996), outro marco histórico da legislação educacional brasileira, determinou que a União deveria se encarregar do processo nacional de avaliação do rendimento escolar em todos os níveis de ensino e do processo de avaliação das IES:

Art. 9. A União incumbir-se-á de:

[...] VI - assegurar **processo nacional de avaliação do rendimento escolar no ensino fundamental, médio e superior**, em colaboração com os sistemas de ensino, objetivando a definição de prioridades e a **melhoria da qualidade do ensino**; [...]

VII - baixar normas gerais sobre cursos de graduação e pós-graduação;

VIII - assegurar **processo nacional de avaliação das instituições de educação superior**, com a cooperação dos sistemas que tiverem responsabilidade sobre este nível de ensino;

IX - autorizar, reconhecer, credenciar, supervisionar e **avaliar**, respectivamente, os cursos das instituições de educação superior e os estabelecimentos do seu sistema de ensino.[...]

Art. 46. A autorização e o reconhecimento de cursos, bem como o credenciamento de instituições de educação superior, terão prazos limitados, sendo renovados, periodicamente, após **processo regular de avaliação**.

§ 1º Após um prazo para saneamento de **deficiências eventualmente identificadas pela avaliação** a que se refere este artigo, haverá **reavaliação, que poderá resultar, conforme o caso, em** desativação de cursos e habilitações, em intervenção na instituição, em suspensão temporária de prerrogativas da autonomia, ou em descredenciamento.
(grifo do autor)

As noções já discutidas anteriormente de fiscalização, controle e responsabilização aparecem nesses artigos da LDB de 1996. Dessa forma, a avaliação pode ter como consequências inclusive a “desativação de cursos e habilitações”, ou a “suspensão temporária de prerrogativas da autonomia”, ou o “descredenciamento”.

Deve-se ainda destacar que as regulamentações propostas na LDB de 1996 acompanhavam tendências nacionais e internacionais. Segundo Bonamino (2002), na década de 1990 aconteceu o ressurgimento da avaliação educacional em larga escala no Brasil. Nessa época, um panorama nacional e internacional destacou a centralidade atribuída a avaliação educacional. Para autora, essa importância é atestada em âmbito

⁵ Ambos os artigos foram revogados posteriormente pela Lei 10.861, de 2004.

nacional pela LDB de 1996 (que estabeleceu a obrigatoriedade de avaliação dos diferentes níveis de ensino) e pela participação brasileira, a partir de 1998, em diversos projetos internacionais de avaliação comparada⁶. Outros autores também destacam a centralidade da avaliação educacional na década de 1990 na agenda internacional. Afonso (2002), por exemplo, ilustra essa questão com um relatório da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) elaborado no ano de 1991. Esse relatório previa que, devido a um crescente interesse pela qualidade do ensino em diferentes países, a avaliação seria um tema amplamente debatido e discutido durante a década de 1990.

Assim, indicamos como a legislação sobre a avaliação educacional acompanhou transformações em um nível mais amplo. A legislação aparece como uma forma de tentar garantir um padrão mínimo de qualidade da educação por meio da avaliação. As diferentes legislações culminaram na atual forma de avaliação do ensino superior, o Sistema Nacional de Avaliação da Educação Superior (SINAES). Por isso, dedicamos nossa próxima seção para a análise do SINAES.

1.4 - O Sistema Nacional de Avaliação da Educação Superior (SINAES)

Como já discutimos anteriormente, o Estado tem procurado controlar a qualidade do ensino das IES públicas e privadas por meio da avaliação. Nesse sentido, a atual política governamental de avaliação do ensino superior continua dentro da lógica do Estado Avaliador.

A forma de avaliação do ensino superior em vigor hoje começou a ser discutida em 2003, ano em que foi criada a chamada Comissão Especial de Avaliação - CEA. Esta comissão tinha como função fornecer subsídios para a reformulação das políticas de avaliação da educação superior e elaborar uma revisão crítica dos métodos e técnicas que vinham sendo utilizados. O texto elaborado pela CEA intitulado “Bases para uma nova proposta de educação superior” foi submetido a amplos debates envolvendo MEC, fórum de reitores, sindicatos, academia, parlamentares e sociedade civil. O resultado da convergência entre o texto e as diversas sugestões encaminhadas foi a lei 10.861 de

⁶ Podemos citar como exemplo o *Programme for International Student Assessment - Pisa* (em português Programa Internacional de Avaliação de Alunos). Ele é um programa de avaliação comparada coordenado internacionalmente pela Organização para Cooperação e Desenvolvimento Econômico - OCDE. O Pisa é aplicado a alunos na faixa etária dos 15 anos. As avaliações acontecem a cada três anos e abrangem as áreas de Leitura, Matemática e Ciências, variando a ênfase em cada uma dessas áreas de acordo com a edição do programa. No Brasil, a coordenação do Pisa é feita pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - Inep (INEP, 2010).

2004, que instituiu o Sistema Nacional de Avaliação da Educação Superior - SINAES (REIS, 2009).

Essa lei consolida legislações anteriores, sendo que já em seu art. 1 faz uma referência à LDB de 1996 ao informar que o SINAES tem como objetivo “assegurar processo nacional de avaliação das instituições de educação superior, dos cursos de graduação e do desempenho acadêmico de seus estudantes, nos termos do art 9º, VI, VIII e IX, da Lei nº 9.394, de 20 de dezembro de 1996”. Desde a criação do SINAES em 2004 até o presente momento, foi editado um grande número de documentos complementares à lei 10.861 de 2004. São alguns exemplos a Portaria 2.051 de 2004, o Decreto 5.773 de 2006 e a Portaria Normativa n.12 de 2008.⁷

A FIG. 1 ilustra o funcionamento atual do SINAES:

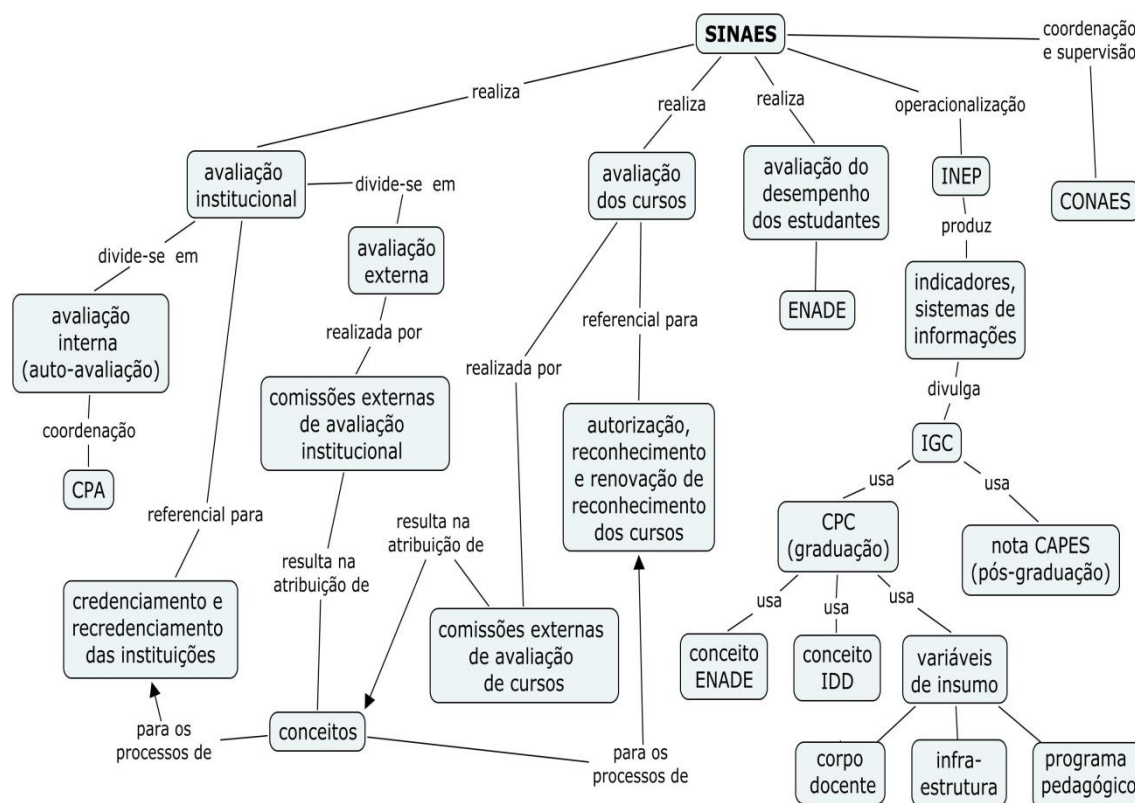


FIGURA 1 - O Sistema Nacional de Avaliação da Educação Superior (SINAES)

Nota sobre as siglas: Comissão Própria de Avaliação (CPA); Exame Nacional de Desempenho dos Estudantes (ENADE); Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP); Comissão Nacional de Avaliação da Educação Superior (CONAES); Índice Geral de Cursos da Instituição de Educação Superior (IGC); Conceito preliminar de curso (CPC); Indicador de Diferença Entre os Desempenhos Observado e Esperado (IDD).

⁷ Segundo Lemos (2008), na hierarquia das leis existem os chamados atos primários (leis) e os chamados atos secundários (resoluções, pareceres), que estão verticalmente abaixo dos primeiros e, geralmente, desempenham o papel de regulamentar os atos primários. É exatamente esse processo que vem acontecendo desde a promulgação da lei 10.861 de 2004.

De acordo com a FIG. 1, o SINAES realiza um conjunto de avaliações. O art. 58 do Decreto 5.773 de 2006 dispõe sobre essa questão:

Art. 58. A avaliação **das instituições** de educação superior, **dos cursos** de graduação e **do desempenho acadêmico de seus estudantes** será realizada no âmbito do SINAES, nos termos da legislação aplicável.

§ 1º O SINAES, a fim de cumprir seus objetivos e atender a suas finalidades constitucionais e legais, compreende os seguintes processos de avaliação institucional:

I - **avaliação interna** das instituições de educação superior;

II - **avaliação externa** das instituições de educação superior;

III - avaliação dos cursos de graduação; e

IV - avaliação do desempenho acadêmico dos estudantes de cursos de graduação.

§ 2º Os processos de avaliação obedecerão ao disposto no art. 2º da **Lei nº 10.861, de 2004**.

(grifo do autor)

Assim, o SINAES possui três componentes principais: avaliação institucional, avaliação dos cursos e avaliação do desempenho dos estudantes. A avaliação institucional se divide em avaliação interna e externa. A avaliação interna (autoavaliação) é coordenada por uma Comissão Própria de Avaliação - CPA. Segundo o art. 11 da lei 10.861 de 2004, toda IES, pública ou privada, “constituirá Comissão Própria de Avaliação - CPA, [...] com as atribuições de condução dos processos de avaliação internos da instituição, de sistematização e de prestação das informações solicitadas pelo INEP [...]”. O mesmo artigo ainda prevê que a CPA precisa ter uma atuação autônoma em relação a órgãos colegiados existentes na IES e assegurar a participação de todos os segmentos da comunidade universitária e também da sociedade civil organizada. É vedada uma composição que privilegie a maioria absoluta de algum dos segmentos. O INEP ainda divulgou em 2004 um documento intitulado “Orientações Gerais para o Roteiro da Auto-Avaliação das Instituições” que tem como público alvo as CPAs, professores, estudantes e técnico-administrativos das IES brasileiras. Esse documento tem orientações e sugestões para o roteiro de avaliação interna que compõe o processo de avaliação institucional. A avaliação externa deve acontecer após o processo de autoavaliação, sendo realizadas por comissões externas de avaliação institucional indicadas pelo INEP. O art. 15 da Portaria 2.051 de 2004 lista todas as informações e documentos examinados por essa comissão externa. São alguns exemplos: o Plano de Desenvolvimento Institucional (PDI); relatórios parciais e finais do processo de autoavaliação; dados do questionário socioeconômico dos estudantes

aplicado no ENADE; dentre outros. A avaliação institucional externa resulta na atribuição de conceitos usados nos processos de credenciamento e credenciamento de instituições.

A avaliação dos cursos de graduação é realizada por comissões externas de avaliação de cursos indicadas pelo INEP. O art. 20 da Portaria 2.051 de 2004 lista todos os aspectos considerados por essa comissão: perfil do corpo docente; instalações físicas; organização didático-pedagógica; desempenho dos estudantes no ENADE; dados do questionário socioeconômico dos estudantes; dados do Censo da Educação Superior e do Cadastro Geral das Instituições e Cursos; outros considerados pertinentes. A avaliação externa dos cursos de graduação resulta na atribuição de conceitos usados nos processos de autorização, reconhecimento e renovação de reconhecimento de cursos. De acordo com o art. 32 da Portaria 2.051 de 2004, a avaliação externa, tanto das instituições quanto dos cursos resulta na “atribuição de conceitos a cada uma e ao conjunto das dimensões avaliadas, numa escala de cinco níveis, sendo os níveis 4 e 5 indicativos de pontos fortes, os níveis 1 e 2 indicativos de pontos fracos e o nível 3 indicativo do mínimo aceitável [...]”.

A avaliação do desempenho dos estudantes é feita por meio do ENADE⁸, que foi instituído pela lei 10.861 de 2004:

Art. 5º A avaliação do desempenho dos estudantes dos cursos de graduação será realizada mediante aplicação do Exame Nacional de Desempenho dos Estudantes - **ENADE**. [...]

§ 2º O ENADE será aplicado periodicamente, admitida a utilização de procedimentos amostrais, aos alunos de todos os cursos de graduação, ao **final do primeiro e do último ano de curso**. [...]

§ 4º A aplicação do ENADE será acompanhada de instrumento destinado a levantar o **perfil dos estudantes**, relevante para a compreensão de seus resultados.

§ 5º O ENADE é **componente curricular obrigatório** dos cursos de graduação, sendo inscrita no histórico escolar do estudante somente a sua situação regular com relação a essa obrigação, atestada pela sua efetiva participação [...]

(grifo do autor)

Como indicado no artigo, participam do ENADE alunos do final do primeiro e do último ano de curso. Esse procedimento tem o objetivo de averiguar o quanto o curso acrescentou aos estudantes, verificando como eles entraram e como estão saindo. Isso foi um ganho comparando com o ENC. Para auxiliar a interpretação dos resultados desse exame, os alunos respondem um questionário socioeconômico. Por fim, merece

⁸ Substituiu o antigo Exame Nacional de Cursos - ENC, da legislação anterior.

destaque o caráter compulsório do ENADE: ele é “componente curricular obrigatório dos cursos de graduação”.

Todo esse processo de avaliação do ensino superior é coordenado e supervisionado pela Comissão Nacional de Avaliação da Educação Superior - CONAES. Esse órgão colegiado de coordenação e supervisão do SINAES foi instituído no art. 6 da lei 10.861 de 2004. Já o art. 7 da mesma lei determina que a CONAES é composta por representantes: do INEP; da CAPES; do Ministério da Educação; dos corpo discente, docente e técnico-administrativo das IES; e membros indicados pelo Ministro da Educação (cidadãos com notório saber e reconhecida competência em avaliação ou gestão da educação superior).

Já a operacionalização do SINAES é feita pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), que produz indicadores e sistemas de informações sobre o ensino superior. Podemos citar como exemplo a divulgação anual do Índice Geral de Cursos da Instituição de Educação Superior - IGC, que foi instituído pela Portaria Normativa n.12 de 2008.

O IGC é um indicador de qualidade das IES que considera em sua composição a qualidade dos cursos de graduação e de pós-graduação. Na graduação, é utilizado o conceito preliminar de curso - CPC. O resultado final é dado em valores contínuos de 0 a 500 e em faixas de 1 a 5. O CPC usa o conceito ENADE, o conceito IDD e as variáveis de insumo (corpo docente, infra-estrutura e programa pedagógico). O Indicador de Diferença Entre os Desempenhos Observado e Esperado - IDD é a diferença entre o desempenho médio do concluinte de um curso e o desempenho médio estimado para os concluintes desse mesmo curso. Ele representa quanto cada curso se destaca da média levando em consideração o perfil dos seus alunos (INEP, 2010).

Segundo Vieira (2008), as avaliações implantadas a partir de 1995, decorrentes do governo de Fernando Henrique Cardoso, reservaram ao Estado o papel de estabelecer regras e de controlar por meio da avaliação o desempenho das IES. A resposta da comunidade acadêmica foi criticar e boicotar duramente o Exame Nacional de Cursos - ENC, sob a crítica de um modelo de avaliação com instrumentos desarticulados e orientados à elaboração de *ranking*, enfatizando os produtos. Quando foi instituída a Comissão Especial de Avaliação - CEA no governo Lula em 2003, o meio acadêmico criou uma expectativa de revisão do sistema de avaliação do ensino superior e de elaboração de propostas novas. No entanto, segundo a autora, o SINAES apresenta uma tensão entre duas concepções de avaliação: por um lado guarda no texto da lei coerência com a concepção de regulação e controle adotada no governo de FHC

no ENC (Estado Avaliador); por outro lado revela, em textos e documentos da CEA e CONAES, outra concepção baseada na formação e emancipação.

Um exemplo da manutenção da concepção de regulação e controle foi a transformação do ENC no ENADE. Quanto à concepção de formação e emancipação, o art. 1 da lei 10.861 de 2004 faz referência a “afirmação da autonomia institucional”. O art. 2 menciona “avaliação institucional interna e externa”, “respeito à identidade e à diversidade de instituições e de cursos”; “participação do corpo discente, docente e técnico-administrativo das instituições de educação superior, e da sociedade civil, por meio de suas representações”. Assim, o incentivo da avaliação interna visa estimular um processo de autoconhecimento da instituição, possuindo certo caráter formativo. O texto também tenta garantir que serão levadas em consideração diversidade e identidade das IES. Portanto, o SINAES busca integrar tanto uma dimensão de regulação quanto uma dimensão formativa de avaliação.

Contudo, o SINAES ainda continua sendo alvo de críticas como: exames nacionais que estabelecem uma classificação (*ranking*) das IES e de seus cursos; a aplicação nas IES de uma lógica empresarial (medição, comparação da produtividade, competitividade); entre outras.

Por fim, consideramos que o SINAES acabou instaurando uma “cultura de avaliação” nas IES: criação de Comissão Própria de Avaliação; os professores são avaliados (inclusive pelos alunos); os estudantes participam do ENADE; o projeto pedagógico e a infra-estrutura da IES são avaliados, entre outros. Tudo isso acaba reverberando dentro da sala de aula. Por exemplo: um determinado professor pode passar a incluir em suas provas questões de interpretação de textos, charges ou tabelas porque o ENADE cobra isso na parte de formação geral. Ou então, esse mesmo docente passa a enfatizar mais determinados conteúdos que estão sendo muito cobrados no ENADE (e os alunos passam a estudar mais esses conteúdos).

Concluimos mais uma vez que entender como os estudantes universitários estão percebendo a avaliação pode nos apontar caminhos para uma melhor compreensão da avaliação que tem sido colocada em prática no ensino superior. Nesse sentido, o próximo capítulo apresenta uma revisão da literatura especificamente sobre as concepções de avaliação dos alunos.

CAPÍTULO 2 - REVISÃO DA LITERATURA

2.1 - Concepções de avaliação

Encontramos na literatura quatro grandes grupos de concepções de avaliação: 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 (responsabilização - *accountability*) (BROWN, 2008; BROWN; HIRSCHFELD, 2008). Os alunos parecem estar conscientes de todas essas quatro concepções e tendem a ter uma orientação direcionada a uma concepção ou uma combinação de várias delas. Nesse sentido, alguns trabalhos sobre o nível de ensino fundamental em diferentes países revelaram concepções variadas de avaliação (ATKINSON, 2003; REAY; WILIAM, 1999). Da mesma forma, no nível de ensino médio e no ensino superior, concepções múltiplas e conflitantes de avaliação ficaram evidentes (CAMARGO, 1997; MONI *et al.*, 2002).

A concepção de melhora enfatiza a expectativa dos alunos de que a avaliação leva a uma aprendizagem melhor, mais eficaz (HARRIS; BROWN; HARNETT, 2009). Pajares e Graham (1998), por exemplo, relataram que alunos de oitava série esperavam um tipo de *feedback* diferente do que eles estavam recebendo dos professores. Eles queriam um *feedback* honesto, compreensível e construtivo sobre como melhorar. Em contraste, os professores ofereciam elogios e sentimentos positivos como respostas. Eles pareciam estar demasiadamente preocupados com o bem estar psicológico dos estudantes e não estavam oferecendo o tipo de *feedback* orientado para a melhora da aprendizagem que os alunos queriam. Dessa forma, existem evidências de que os alunos concebem a avaliação como tendo um propósito de melhora.

A concepção de que a avaliação tem um impacto emocionalmente positivo nos alunos enfatiza as respostas emocionais dos estudantes com relação a formatos diferentes de avaliação (BROWN; IRVING; PETERSON; HIRSCHFELD, 2009). Existem evidências de que os alunos simplesmente preferem qualquer sistema de avaliação que eles estejam experienciando, independente dos méritos ou deficiências desse sistema (BLAIKIE; SCHONAU; STEERS, 2004). Ou seja: os estudantes talvez simplesmente reproduzam e adotem os valores, crenças e preferências dos seus professores (ZEIDNER, 1992). Deve-se ainda destacar que uma boa parte da literatura sobre as experiências ou percepções de avaliação dos alunos está centrada nas atitudes dos alunos com relação a tipos diferentes de avaliação (BROWN, 2008). Respostas emocionais positivas como a motivação dos colegas de sala de aula e a avaliação como

uma experiência agradável estão incluídas nessa concepção de avaliação. Assim, estudos como esses demonstram que alguns alunos descrevem respostas emocionais positivas com relação à avaliação.

A concepção de irrelevância enfatiza a percepção de que a avaliação é negativa, ruim ou injusta. Existem evidências de que os alunos diferem quanto ao grau de percepção negativa da avaliação. A maioria percebe a avaliação de forma cada vez mais negativa na medida em que vai avançando no sistema educacional (BROWN; IRVING; PETERSON; HIRSCHFELD, 2009). Por exemplo, Moni, van Kraayenoord e Baker (2002) relataram que as atitudes com relação à avaliação de 54 estudantes australianos do primeiro ano do ensino médio se tornaram bem mais negativas quando comparadas com o ensino fundamental por dois motivos: o maior número de avaliações do ensino médio; a percepção de uma subjetividade nas decisões de avaliação do professor. Já alunos do ensino superior perceberam a avaliação como arbitrária, irrelevante e inexata, sendo simplesmente um processo necessário para dar notas. As razões apontadas pelos alunos foram a percepção de subjetividade e falta de profissionalismo na atribuição de notas nas tarefas ou exames (DUFFIELD; SPENCER, 2002; HAWES, 2002). Assim, os alunos se mostram bem sensíveis quanto à avaliação que eles percebem como injusta, ruim ou irrelevante para eles.

A concepção de responsabilização enfatiza a ideia de que a avaliação torna os alunos responsáveis pelo seu aprendizado, por meio de consequências como notas ou certificados (HARRIS; BROWN; HARNETT, 2009). Além disso, nessa concepção a avaliação também pode se mostrar relacionada com fatores externos ao indivíduo, por exemplo, a escola, os pais ou o futuro profissional. Ou seja, também pode se relacionar com um *locus* de controle externo (BROWN; PETERSON; IRVING, 2009). A existência dessa concepção pode ser inferida pelas respostas que os alunos dão para avaliações que tem consequências importantes para eles⁹. Wheelock, Bebell e Haney (2000), por exemplo, descobriram que alunos americanos da quarta, oitava e décima séries estavam conscientes de que os resultados dos testes tinham sérias consequências para eles incluindo questões como retenção ou formatura (especialmente os estudantes do ensino médio). Reay e Wiliam (1999) relataram um estudo onde foi desenhada uma avaliação nacional para monitorar a qualidade e eficácia das escolas britânicas. Mesmo

⁹ O termo em inglês é *high-stakes test* e pode ser definido como um teste que tem consequências importantes para o examinando. Se o examinando passar no teste, ele pode receber benefícios significativos como, por exemplo, um diploma, uma bolsa ou a aprovação em um concurso. Um exemplo desse tipo de teste no Brasil é o vestibular.

nessa situação, alunos do ensino fundamental trataram as notas baixas como uma medida deles mesmos ao invés da escola ou do professor. Dessa forma, embora alguns alunos possam estar conscientes de que a avaliação pode melhorar o ensino/aprendizagem ou tornar as escolas responsáveis, estudos como esses fornecem evidências de que os alunos veem a avaliação como algo que os torna responsáveis pelo seu aprendizado (BROWN, 2008).

Como afirmamos anteriormente, os alunos podem ter uma orientação direcionada a uma concepção ou uma combinação de diversas concepções. Uma fonte importante para entender essa questão e determinar as características das concepções são as pesquisas sobre conceitos. Segundo Brown (2008), um aspecto central do debate nessa área é se os nossos conceitos são sistemáticos, organizados e consistentes ou se são desorganizados e inconsistentes. Acreditamos que a inconsistência dos conceitos é mais provável do que a consistência. Sambell e McDowell (1998), por exemplo, descreveram em três estudos de caso como a mesma atividade de avaliação foi percebida de maneiras opostas pelos alunos. Além disso, um mesmo aluno pode relatar que a avaliação serve para melhorar o aprendizado e que ela é injusta e algo a ser ignorado (concepções de avaliação que parecem ser opostas e conflitantes). A própria natureza do fenômeno educacional experienciado por alunos e professores parece contribuir para isso. Para Clark e Peterson (1986), os professores teriam uma maneira de pensar inconsistente e incompleta porque a maioria dos problemas que enfrentam são grandes dilemas sem uma solução fácil e imediata. Ao pensarmos isso na realidade brasileira, a situação se torna ainda mais saliente. Portanto, acreditamos que a noção de que as concepções são sistemáticas e organizadas podem não ser adequadas para dar conta da complexidade do pensamento dos alunos e professores.

2.1.1 - A questão da terminologia nos estudos sobre as concepções de avaliação

Nas pesquisas sobre o que pensam alunos e professores, uma enorme variedade de termos tem sido usada para se referir às representações mentais. Encontramos termos como: concepções, conceitos, crenças, imagens, metáforas, percepções, orientações, perspectivas, categorias, construtos, conhecimentos, culturas, repertórios, teorias, representações, dentre outros. Essa variedade de formas pode ser encontrada por vezes dentro de um mesmo texto. Destacamos ainda que alguns desses termos são utilizados como sinônimos em determinados trabalhos. Isso acaba dificultando a comparação e compreensão de pesquisas das áreas educacional e psicológica. No entanto, os termos utilizados nas pesquisas sobre concepções de avaliação acabam chegando a um mesmo

resultado: informar a maneira como as pessoas percebem, avaliam e agem com relação ao fenômeno da avaliação. Existe um consenso entre pesquisadores que investigam concepções de que essas diferenças são mais semânticas do que substanciais (ETHELL; MCMENIMAN, 2002).

Dessa forma, um aspecto recorrente na área de Ciências Humanas se faz presente nos estudos sobre as concepções de avaliação: a questão da terminologia. Assim como outros termos da área de pesquisa social, o termo “concepção” carrega uma variedade de significados (QUADRO 1). Além disso, como já apontamos anteriormente, muitas outras palavras têm sido usadas para se referir às concepções de alunos e professores.

QUADRO 1
Definições de concepção

Representações mentais, sistemas complexos de explicação (WHITE, 1994).

[...] rede complexa de ideias, conceitos, representações e, inclusive, preconceitos - em seu sentido valorativo (PERDIGÃO, 2002, p. 268).

Psicologia: operação pela qual o espírito forma um conceito.

Psicologia: fase primeira do ato voluntário (concepção do objetivo) (RUSS, 1994, p. 45).

Operação para a qual o espírito constrói, sem necessariamente apelar para dados experimentais, um conceito ou ideia geral. É também chamada hoje em dia de conceitualização. Na psicologia clássica, é o primeiro momento do ato voluntário: aí o espírito concebe a meta a ser alcançada (DUROZOI; ROUSSEL, 1996, p. 98).

A partir do QUADRO 1, destacamos dois grandes sentidos da definição de concepção. Um primeiro mais geral com definições como “sistemas complexos de explicação” e “rede complexa de ideias, conceitos, representações e preconceitos”. Assim, as concepções podem informar a maneira como as pessoas percebem, avaliam e agem com relação a um determinado fenômeno. O segundo sentido mais específico aparece nas definições como a “operação de construção de conceitos”. Dessa forma, as nossas concepções envolveriam um processo de formação de conceitos. Para uma discussão mais aprofundada sobre a definição de concepção, ver Brown (2008) e Lopes (2007).

Concluimos apontando que, nesse trabalho, o termo concepção é usado para se referir a “sistemas complexos de explicação”, uma “rede complexa de ideias, conceitos, representações e preconceitos”.

2.2 - As concepções de avaliação dos alunos universitários brasileiros: uma revisão de estudos

Struyven *et al.* (2005) realizaram uma extensa revisão da literatura com estudos que abordavam as percepções de alunos sobre avaliação no ensino superior. O método utilizado incluiu uma busca *online* de trabalhos nas bases de dados *Educational Resources Information Center* (ERIC), *Web of Science* e *PsychINFO* envolvendo os anos de 1980 até 2002. As palavras chave “percepção do aluno” e “avaliação” foram combinadas, resultando em 508 artigos nas bases de dados ERIC e *PsycINFO* e 37 na *Web of Science*. Quando a pesquisa foi limitada com a palavra chave adicional “ensino superior”, restaram 171 artigos nas bases ERIC e *PsycINFO* e 10 na *Web of Science*. Considerando outros critérios de seleção, no final foram analisados 35 trabalhos. Os critérios para a seleção incluíram pesquisas quantitativas e qualitativas. Um importante argumento apresentado pelos autores foi que as percepções dos alunos sobre a avaliação influenciam significativamente a abordagem deles em relação à aprendizagem e ao estudo. Os trabalhos revisados evidenciaram que as percepções dos alunos sobre avaliação e as suas abordagens em relação à aprendizagem estão fortemente ligadas. Além disso, as percepções dos alunos sobre a avaliação também se mostraram fortemente associadas com o formato da avaliação e os métodos utilizados pelo professor (métodos de avaliação tradicionais ou alternativos) e com a percepção dos alunos sobre a justiça da avaliação. Todas essas influências podem ser tanto positivas quanto negativas. Portanto, essa revisão da literatura indicou que as percepções dos alunos e os julgamentos que eles fazem das características das avaliações (formato de avaliação, justiça) afetam tanto positivamente quanto negativamente a abordagem dos estudantes em relação à aprendizagem e ao estudo no nível de ensino superior. Essa revisão, assim como outras revisões internacionais, não incluiu pesquisas realizadas no Brasil. Consequentemente, realizamos uma revisão de estudos sobre as concepções de avaliação dos alunos universitários brasileiros para tornar disponíveis informações sobre essa temática tanto nacional quanto internacionalmente.

Vasconcellos, Oliveira e Berbel (2006), por exemplo, analisaram as concepções de avaliação de alunos e professores de diversos cursos de uma universidade pública. Essa pesquisa se iniciou com uma investigação junto aos alunos sobre as práticas avaliativas consideradas como positivas e negativas vivenciadas por eles. No total, 428 estudantes pertencentes a 14 cursos de licenciaturas responderam um questionário. Esse instrumento, constituído por duas perguntas abertas, solicitava aos alunos que relatassem as experiências avaliativas que os tivessem marcado positiva ou

negativamente. Posteriormente, foram selecionadas as experiências positivas. Em alguns casos, os estudantes indicaram também o nome dos professores responsáveis por essas práticas avaliativas positivas. No total, 48 docentes foram indicados. A amostra de professores foi constituída por 41 licenciados e sete bacharéis. Utilizando um método qualitativo, foram realizadas entrevistas semi-estruturadas com os docentes. Foi realizada com esse material uma análise de conteúdo, extraindo das respostas às entrevistas elementos convergentes para alguns focos tomados para reflexão. Nos resultados, os autores destacam três aspectos que representam temáticas para reflexões a respeito de práticas avaliativas: a) experiências pessoais - os docentes atuam também por motivos individuais e de acordo com suas experiências pregressas. Assim, os professores têm uma tendência para repetir práticas de avaliação passadas. Na realidade, experiências de avaliação que eles tenham vivido como estudantes tiveram um papel significativo nas práticas de avaliação dos professores; b) reflexões geradas pelo processo de autoavaliação - a autoavaliação é apontada como um exercício que possibilita ao professor refletir sobre sua prática, revelando motivações e intenções. Eles monitoraram as reações dos alunos a várias práticas de avaliação e modificaram as suas práticas levando em conta esse *feedback*; c) valores e concepções que os docentes possuem - embora muitos procedimentos de avaliação estejam disponíveis, os professores enfatizaram que o aspecto mais importante era a intenção do avaliador quanto ao uso de um tipo específico de avaliação. Qualquer que seja o tipo de prática selecionada tem que resultar em uma aprendizagem efetiva para os estudantes. Essa concepção orientada para a melhoria era o paradigma dominante para professores bem avaliados. No entanto, fica aparente que a ênfase neste estudo estava nas visões dos professores.

Assim, realizamos uma busca da literatura brasileira. Bancos de dados brasileiros (por exemplo, Scielo e Pepsic), e a Biblioteca Digital brasileira de Teses e Dissertações foram pesquisados na internet. Nós combinamos os seguintes descritores usando tanto o singular quanto o plural: concepção, percepção, representação, crença, aluno, estudante e avaliação¹⁰.

Também pesquisamos bancos de dados internacionais (*Academic Search Premier, Mental Measurements Yearbook, Professional Development Collection, PsycINFO, Psychology and Behavioral Sciences Collection, and Research Starters - Education*) adicionando a palavra Brasil (por exemplo, “concepção aluno” e “avaliação”

¹⁰ Ver o item “2.1.1 - A questão da terminologia nos estudos sobre as concepções de avaliação”, que explica a necessidade do uso de todos esses descritores.

e “Brasil”). Documentos relevantes foram buscados e selecionados. Além disso, às vezes encontramos somente os resumos dos estudos. Assim, esses trabalhos não foram incluídos nessa revisão da literatura porque não atendiam o nosso critério: fornecer informações suficientes para descrever e analisar os resultados da pesquisa.

Nossa revisão da literatura sobre as concepções de avaliação dos alunos do ensino superior é classificada pelo número de concepções informado (poucas ou múltiplas concepções). A revisão da literatura também apresenta estudos que discutem as perspectivas dos alunos sobre tipos de avaliação.

2.2.1 - Estudos com poucas concepções

A maioria dos estudos identificou somente uma ou duas concepções entre os alunos. Camargo (1997), por exemplo, usou análise do discurso para investigar as concepções de avaliação dos alunos analisando a prática da avaliação escolar no ensino superior. A autora, utilizando uma teoria da produção do discurso bakhtiniana e a teoria das representações sociais, explicitou as representações e os sentidos que os estudantes construíram mediante suas experiências em avaliação. Ela analisou relatos escritos de 390 universitários que cursaram a disciplina Avaliação e Medidas Educacionais de um curso de Pedagogia, formandos no período compreendido entre 1991 e 1994. Os resultados indicaram uma predominância de experiências negativas (80%) em relação à avaliação na história escolar dos alunos. No discurso dos estudantes apareceram frequentemente os seguintes aspectos: referências ao uso da avaliação como instrumento de disciplina e dominação; imposição externa de normas e ordens por parte do professor e ausência de adesão interna por parte do aluno; tolhimento à construção da identidade intelectual do aluno; problemas na comunicação pedagógica. Os episódios positivos revelados nos relatos dos alunos mostraram: prática da avaliação com um caráter contínuo e processual; instrumentos de avaliação diversificados; exigência de habilidades intelectuais complexas e não a memorização; professor competente na didática e no conteúdo ministrado; professor interessado na aprendizagem do aluno, que dá retorno sobre os trabalhos realizados e orienta a correção dos erros. Assim, a autora considera que o eixo organizador das representações positivas da avaliação evidencia uma qualificação do aluno como um sujeito cognoscente e interativo. Tanto a concepção de irrelevância quanto a concepção de melhora ficaram aparentes neste estudo.

Neves (2002) investigou as representações de avaliação da oralidade na aprendizagem de inglês como Língua Estrangeira (LE) em uma universidade pública. A

autora usou uma abordagem transdisciplinar, agregando contribuições teóricas da análise do discurso (por exemplo, Michel Foucault, Michel Pêcheux), linguística aplicada e psicanálise. A amostra consistiu em 15 estudantes e 4 professores que responderam um questionário aberto com 32 perguntas (por exemplo: *O que dizem os nativos falantes do inglês sobre o seu desempenho nessa língua? Você tem chances de usar a língua estrangeira fora da sala de aula? Como?*). Práticas avaliativas tradicionais na aprendizagem de inglês como língua estrangeira em cursos de graduação como fluência e pronúncia, competência comunicativa/proficiência, e avaliação dos colegas foram examinadas. As representações de avaliação de alunos e professores foram classificadas principalmente em duas categorias (formações discursivas): inclusiva e excludente. A primeira categoria se caracteriza pela aceitação da falta e pela inclusão do progresso do aluno. Além disso, associa a avaliação com procedimentos inclusivos como a aceitação de erros. A segunda caracteriza-se por um imaginário de perfeição (o desempenho do aluno precisa ser perfeito) e associa a avaliação com procedimentos excludentes como a retenção e padrões de desempenho. Assim, a concepções de melhora e de responsabilização apareceram nesta pesquisa.

Cacione e Souza (2005) realizaram um estudo de caso na Universidade Estadual de Londrina (UEL). A amostra do estudo foi de 10 estudantes do terceiro e quarto ano do curso de licenciatura em música da UEL. As autoras utilizaram como procedimentos metodológicos questionários, entrevistas semi-estruturadas e análise de documentos. Os resultados foram classificados em três categorias: conceito de avaliação da aprendizagem; modalidades e funções da avaliação da aprendizagem; técnicas e instrumentos utilizados para a avaliação da aprendizagem. A maioria dos alunos demonstrou possuir uma concepção de avaliação da aprendizagem ligada a ideias conservadoras. Também foram relatadas práticas avaliativas que efetuam comparações entre sujeitos e que muitas vezes desconsideram outros objetivos traçados. Assim, os estudantes associaram a avaliação com instrumentos tradicionais, indicando que as técnicas e instrumentos avaliativos não são usados para acompanhar o processo ensino-aprendizagem. Eles são usados apenas para obter notas e possibilitar a promoção no final do ano. Uma minoria dos alunos indicou uma vivência ou proposta de prática avaliativa mais preocupada com o processo e não simplesmente com o produto. Entretanto, mesmo nesses casos, a percepção dos alunos indica uma inconsistência entre fundamentação teórica e prática. Portanto, as concepções de avaliação estavam profundamente ligadas à concepção de responsabilização (a avaliação para aprovar ou reter). Somente alguns estudantes relataram experiências ligadas à avaliação formativa.

Pereira (2006) realizou um estudo de caso em um curso de Pedagogia de uma instituição pública federal do Estado de Minas Gerais. A autora utilizou como procedimentos metodológicos: pesquisa documental (análise do Projeto Político Pedagógico do curso) e um questionário composto por 14 questões objetivas e seis abertas. Foram selecionados para a amostra somente alunos do final do curso (4º ano). No total, 57 estudantes responderam ao questionário. As perguntas abertas desse instrumento buscavam identificar como os alunos apreendem a avaliação da aprendizagem ensinada no curso de formação de professores. As perguntas envolviam os seguintes aspectos: a caracterização de uma boa avaliação; os aprendizados significativos teóricos e práticos sobre avaliação para a atuação profissional; os procedimentos mais comuns de avaliação do curso; a indicação da melhor e pior forma de avaliação vivida no decorrer do curso; o tratamento dado pelos professores aos erros e aos acertos. Na análise dos resultados, a autora destaca que os alunos esperavam métodos de avaliação bastante diferentes daqueles que estavam sendo usados pelos seus professores. Os estudantes queriam mais do que as práticas de avaliação tradicionais e mostraram uma consciência crítica sobre questões teóricas e práticas. Dessa forma, a maioria dos sujeitos relatou a contradição entre a teoria e a prática avaliativa como um dos grandes problemas da formação. Portanto, eles perceberam uma distância entre as recomendações teóricas e as práticas de ensino reais da universidade. Neste estudo, parece que a avaliação é concebida teoricamente como melhora. Porém, a prática revela uma ênfase na concepção de responsabilização.

Mezzaroba (2000), em um estudo de caso do Curso de Farmácia (Habilitação em Bioquímica) da Universidade Estadual de Londrina (UEL), analisou as concepções de avaliação da aprendizagem de alunos e professores. A autora identificou as concepções e finalidades da avaliação, os critérios para elaboração e correção dos instrumentos de medidas mais utilizados nas disciplinas do ciclo profissionalizante e como os professores relacionam os objetivos de ensino e a avaliação da aprendizagem. Participaram da pesquisa 26 alunos da quinta série do curso e 26 docentes do Departamento de Patologia Aplicada, Legislação e Deontologia do Centro de Ciências da Saúde que atuam em oito disciplinas de Farmácia e Bioquímica da UEL. Os instrumentos utilizados no trabalho foram questionários abertos aplicados a alunos e professores e entrevistas semi-estruturadas. A análise qualitativa das informações obtidas foi feita por meio da técnica de análise de conteúdo proposta por Bardin. A autora classificou os resultados encontrados nos questionários e entrevistas em diversas categorias. Considerando os alunos, a homogeneidade das respostas dadas permitiu o

agrupamento em uma única categoria: verificar o aprendizado/desempenho do aluno, isto é, medir o conhecimento assimilado ou adquirido pelo aluno. Além disso, os estudantes demonstraram sentimentos negativos com relação à avaliação (medo, ansiedade, submissão e passividade). Três categorias demonstraram, em ordem de importância, a concepção de avaliação para os professores: mensurar ou quantificar o conhecimento do aluno; verificar o alcance dos objetivos de ensino; mensurar ou verificar a metodologia de ensino ou o desempenho do docente. Tanto uma forte concepção de responsabilização quanto a concepção de irrelevância ficaram evidentes nesta pesquisa.

Pellisson (2007) investigou as percepções de alunos e professores sobre a avaliação de inglês como Língua Estrangeira (LE), assim como as implicações dessas percepções para a prática pedagógica. A amostra consistiu em 75 estudantes e duas professoras de um curso de Letras de uma universidade privada. Foram utilizados entrevistas e questionários semi-estruturados com as professoras e com os alunos, que estavam cursando o início, meio e final do curso (3º, 5º e 8º períodos). Também foram realizadas observações de aulas com anotações de campo. Segundo a autora, a análise dos dados mostrou que existem muitas convergências entre as percepções das professoras e dos alunos. Por exemplo, ambos os grupos tiveram uma percepção da avaliação como algo processual, contínuo e formativo. No entanto, um exame da sala de aula mostrou que o que realmente é praticado é uma avaliação somativa baseada no produto. Na realidade, os alunos têm pouco contato com a temática da avaliação durante o curso, indicando que a graduação não propicia um conhecimento teórico com relação a esse tema. Além disso, estudantes demonstraram percepções negativas conectando a avaliação com nervosismo, burocracia, controle, punição, poder, medo e pressão. Os resultados da pesquisa ainda apontam para uma deficiência na formação do futuro professor de Inglês, inclusive quanto ao processo de avaliação. Nesta pesquisa, parece haver uma concepção de melhora (teórica) na percepção dos estudantes. No entanto, a prática revela uma ênfase nas concepções de responsabilização e irrelevância.

Gesser (1996) investigou as percepções de avaliação institucional de membros de uma instituição de ensino superior privada de Santa Catarina. A amostra foi composta por 63 pessoas, incluindo estudantes pertencentes a cinco cursos de graduação, professores, diretores e funcionários. Os procedimentos metodológicos adotados foram a análise ambiental e entrevistas individuais e grupais. Resultados de uma análise de conteúdo mostraram que as percepções de avaliação institucional eram frequentemente associadas com: a importância do clima institucional, controle,

diagnóstico, processo de tomada de decisões, processo coletivo-democrático inovador e aspectos do ambiente institucional (por exemplo, o desejo de busca de uma melhor qualidade educacional). Este estudo parece indicar tanto a concepção de melhora quanto a concepção de responsabilização (institucional).

2.2.2 - Estudos com concepções múltiplas

Encontramos poucos estudos que examinaram concepções múltiplas simultaneamente. Leite *et al.* (2006) aplicaram um questionário aberto em uma universidade Portuguesa e duas universidades públicas Brasileiras para investigar as percepções dos alunos sobre avaliações administradas para estabelecer a qualidade da educação universitária. Participaram 466 estudantes de 55 cursos (amostra aleatória). Os alunos responderam a duas questões: a) Qual é a sua opinião sobre a avaliação das universidades? e b) você acha que a avaliação da universidade produz ou produzirá melhorias no seu curso? Após um processo rigoroso de codificação qualitativa, os autores identificaram sete temas principais: (1) avaliação e professores; (2) avaliação construtiva; (3) o controle das universidades e regulação pelo estado; (4) validação da avaliação e legitimidade; (5) responsabilização (*accountability*); (6) discrepância e comparação entre universidades; e (7) segmentação/fragmentação da avaliação. Os alunos perceberam a avaliação como um exercício legítimo, resultante de uma decisão política, conduzindo a uma melhoria da qualidade das universidades. Eles também reconheceram que a avaliação está associada com comparações institucionais, mecanismos de controle, regulação, monitoramento e possibilidade de padronização. Além disso, os estudantes tiveram percepções múltiplas sobre a dimensão de ensino. Eles consideraram que a avaliação das universidades poderia ser uma fonte de *feedback* para os professores e poderia melhorar tanto os métodos de ensino quanto o sucesso acadêmico dos estudantes. Porém, os alunos também perceberam que a avaliação não produziu efeitos positivos no desempenho dos professores, especialmente sem um sistema de punição. Assim, esta pesquisa identifica claramente inter-relações complexas entre as concepções de melhora, responsabilização (da universidade) e irrelevância.

Vieira (2006) usou a teoria das Representações Sociais para investigar as representações de 178 estudantes de 8 cursos de licenciaturas de uma universidade em Uberaba - MG sobre o portfólio como instrumento avaliativo. Um primeiro questionário, composto por 19 questões de escolha múltipla e uma pergunta aberta (*escreva cinco palavras que lhe vêm à mente ao ler a palavra em destaque: **Portfólio***), buscou coletar principalmente dados demográficos sobre os alunos como sexo, idade e

nível socioeconômico. No caso da questão aberta, os alunos também eram instruídos a escolher a palavra mais importante e justificar a resposta. Um segundo questionário tinha nove questões de múltipla escolha e três perguntas abertas (*a partir de sua experiência, explique o que é portfólio; se um conhecido seu, estudante de outra instituição, e que não conhece o portfólio, perguntar-lhe o que você acha desse instrumento de avaliação, o que responderia para ele?; caso você queira apontar mais alguma consideração sobre o portfólio, aproveite o espaço abaixo*). Esse instrumento tinha como objetivo identificar as representações sociais dos alunos sobre o portfólio como instrumento de avaliação da aprendizagem. Na análise dos dados, a autora identificou três grupos de representações: 1) aprovação da avaliação realizada por meio do portfólio: a maioria dos estudantes aprova e considera esse instrumento como uma maneira de minimizar avaliações punitivas e excludentes; 2) desaprovação da avaliação realizada por meio do portfólio: uma minoria de alunos não acredita nessa forma de avaliar; 3) aprovação do portfólio, mas com ressalvas: são alunos que, apesar de apoiarem o instrumento, apresentam algumas restrições. Assim, os resultados da pesquisa apontam que os alunos compreendem as intenções do portfólio e aceitam seus propósitos. No entanto, a autora ainda destaca a necessidade de argumentar com os estudantes no sentido de favorecer a aceitação do portfólio. Ou seja: alguns alunos ainda precisam ser convencidos sobre a validade e utilidade desse instrumento avaliativo, inclusive sobre a possibilidade de utilização futura quando esses estudantes desenvolverem sua prática como docentes. Estes resultados sugerem que os alunos estavam atentos às concepções de responsabilização, irrelevância e melhora (embora alguns estudantes ainda precisassem ser persuadidos de que o portfólio é uma estratégia de avaliação legítima).

2.2.3 - Tipos de atividades avaliativas

Existem alguns estudos brasileiros com relação às percepções dos alunos sobre os tipos de atividades avaliativas. Pellisson (2007), por exemplo, realizou uma pesquisa com 75 estudantes de um curso de Letras de uma universidade privada. A autora utilizou entrevistas e questionários para fazer diversas perguntas sobre avaliação, dentre elas: 1) *Dê exemplos de exercícios que o professor usa nas provas* e 2) *Quais os instrumentos de avaliação que seu professor usou neste semestre?* Nas provas, nenhum formato de questão dominou: produzir diálogos representou pouco mais da metade das respostas (51%), enquanto preencher lacunas (49%), interpretação de textos (47%) e questões de múltipla escolha ou dissertativa (45%) representaram cada uma quase a metade das

respostas. No entanto, atividades avaliativas diferentes de provas eram poucas e infrequentes: produção de textos (20%), seminários (7%), resenhas (5%) e debates (3%). Essas atividades tendem a ser formais, mas já permitem algumas práticas mais interativas e informais.

Pereira (2006) também perguntou para 57 estudantes de Pedagogia de uma universidade pública de Minas Gerais sobre os procedimentos de avaliação mais comuns durante o curso. Três tipos de avaliação dominaram: trabalhos (em grupo, individuais, em sala, expositivos, escritos, resumos, fichamentos, esquemas e resenhas) (93%); provas (39%) e seminários (30%). Por outro lado, a autoavaliação recebeu a porcentagem mais baixa (1%). Portanto, as práticas avaliativas eram largamente formais, embora houvesse espaço para práticas informais.

Esses dois estudos sobre o nível superior mostraram uma ênfase comum em métodos de avaliação convencionais e baixo uso de avaliação interativo-informal como a autoavaliação. Em contraste, em duas escolas de ensino fundamental do estado do Maranhão, 14 professores responderam uma pergunta sobre os tipos de atividades avaliativas usadas com mais frequência com os alunos (MACHADO, 2006). Todos os professores informaram o uso de duas atividades avaliativas (trabalho em grupo e observação diária do desempenho dos alunos em sala de aula). Somente a metade dos professores informou o uso de testes ou provas. Porém, as notas eram extraídas em sua maioria de testes e provas (práticas avaliativas convencionais). Já as práticas mais usadas de avaliação informal (trabalhos e observação) desempenharam um papel pequeno na elaboração das notas dos alunos.

Portanto, parece que as práticas avaliativas formais dominam a sala de aula e as práticas universitárias no Brasil. Embora atividades informais possam acontecer, elas parecem ter pouco peso na elaboração das notas.

2.2.4 - Uma análise sobre a revisão de estudos brasileiros

Em relação à revisão da literatura sobre as concepções de avaliação de alunos universitários brasileiros, destacamos alguns pontos importantes:

- a) Mesmo que poucos estudos tenham conseguido captar e inter-relacionar concepções múltiplas de avaliação dos alunos, ficou evidente a partir desta revisão que uma gama extensa de concepções de avaliação foi identificada nas experiências e pensamentos de estudantes universitários brasileiros. Estão incluídas aqui a consciência de responsabilização do aluno, responsabilização da escola, melhora e irrelevância. A concepção de responsabilização aparece oito

vezes nos estudos, a de melhora sete vezes e a de irrelevância cinco vezes. Alertamos para o fato de que a concepção de melhora aparece algumas vezes como uma “situação ideal” ou um “posicionamento teórico” e não prático dos envolvidos no processo. Assim, a concepção dominante parece ser uma reação emocional negativa (irrelevância) com relação à responsabilização do aluno. De forma secundária aparece também uma concepção de avaliação para melhora.

- b) Existem evidências de que formas tradicionais de práticas avaliativas (ex: provas) são mais fortemente associadas com o propósito de responsabilização, enquanto tipos de avaliação mais alternativos e informais (ex: autoavaliação e avaliação feita pelos colegas) parecem estar relacionados com a concepção de melhora. Por fim, existem também evidências de que as práticas avaliativas formais dominam a sala de aula e as práticas universitárias no Brasil.
- c) Considerando especificamente a revisão da literatura que fizemos sobre as concepções de avaliação dos alunos, constatamos uma carência de publicações de trabalhos da área em revistas científicas. Em primeiro lugar, ainda não existe uma cultura no Brasil quanto à publicação de trabalhos do tipo “estado da arte”, mesmo que em grande parte as dissertações de mestrado e as teses de doutorado apresentem revisões da literatura. Entendemos que a publicação desses estudos facilitaria muito a divulgação da produção científica da área. Além disso, durante a realização dessa revisão, muitos trabalhos foram encontrados somente no formato de dissertações ou teses. Por fim, os sistemas de busca e armazenamento de trabalhos no Brasil ainda apresentam muitos problemas. Como às vezes encontramos somente os resumos dos estudos, esses trabalhos não foram incluídos na revisão da literatura. A existência de iniciativas como a “Biblioteca Digital Brasileira de Teses e Dissertações” já representa um avanço. Mas mesmo nesse caso infelizmente nem todas as universidades disponibilizam ou atualizam seus dados nesse site. Destacamos ainda outras dificuldades adicionais como o fato de algumas universidades não liberarem o acesso aos trabalhos na íntegra.
- d) Percebemos na revisão da literatura sobre as concepções de avaliação uma predominância de métodos de pesquisa qualitativos. Entendemos que isso reflete no Brasil não só a produção científica nesse campo de estudos, mas toda a produção acadêmica brasileira. Além disso, as poucas pesquisas encontradas na revisão que apresentavam métodos quantitativos utilizavam apenas métodos de análise muito simples como frequências/porcentagens ou no máximo o uso de

estatística descritiva, demonstrando um descompasso com a literatura internacional da área. Assim, apontamos como uma possível direção para as pesquisas sobre as concepções de avaliação de alunos no Brasil o uso de métodos mais robustos de análise de dados, principalmente de estatística multivariada (esse argumento também é válido para outras áreas da pesquisa educacional). Os pesquisadores deveriam dar mais atenção ao uso de técnicas como Análise Fatorial (Exploratória e Confirmatória), Escalonamento Multidimensional, Modelagem de Equações Estruturais, Teoria de Resposta ao Item, dentre outros métodos. A seguir, discutimos de forma um pouco mais aprofundada a questão dos métodos quantitativos. Essa discussão é importante, pois o uso de métodos quantitativos na pesquisa educacional ainda é incipiente.

Primeiramente, destacamos que as pesquisas baseadas em abordagens quantitativas têm sido em grande parte ignoradas por estudiosos da área educacional. Alguns entendimentos equivocados podem contribuir para essa resistência entre os pesquisadores. Segundo Gatti (2004), é comum encontrar a afirmação de que até meados do século passado predominavam no Brasil estudos de natureza quantitativa, que foram batizados inadequadamente de positivistas ou tecnicistas. Essa afirmação expressa o fato de que durante décadas o uso de métodos de pesquisa quantitativos foi visto por boa parte da comunidade acadêmica brasileira como um resquício do positivismo ou simplesmente uma metodologia acrítica e desprovida de profundidade. Atualmente, a principal crítica se refere a uma postura política neoliberal. Afonso (2002), por exemplo, afirma que a teoria da avaliação, que vinha evoluindo com base em epistemologias anti-positivistas e pluralistas, sofre hoje um novo viés positivista. Isso ficaria evidenciado pela adoção de uma política neoliberal baseada em indicadores mensuráveis.

Destacamos que a resistência com relação ao uso de métodos quantitativos mostra um descompasso com a literatura de metodologia científica, já que há algum tempo se reconhece a complementaridade entre os métodos quantitativos e qualitativos. Na literatura, fica evidenciada a necessidade de adequação do método de acordo com o objetivo da pesquisa. Dessa forma, não existem métodos de pesquisa que sejam bons ou ruins em si mesmos. Nessa direção, Minayo e Sanches (1993) apontam que não devemos construir um continuum entre a pesquisa quantitativa e a qualitativa, tendo num polo a objetividade e no outro a subjetividade. Assim, nenhuma pesquisa é puramente quantitativa. A interpretação das análises estatísticas envolve teoria e

também um processo subjetivo. Já Gatti (2004) faz as seguintes colocações: uma boa análise irá depender de boas perguntas feitas pelo pesquisador; dados como tabelas em si não dizem nada, pois o significado dos resultados é dado pelo pesquisador em função de sua teoria; deve-se conhecer o contexto em que os dados foram produzidos e a área em que os problemas estudados se situam. A autora ainda destaca que os métodos quantitativos podem ser muito úteis na compreensão de vários problemas educacionais, lembrando que a combinação com métodos qualitativos pode enriquecer a compreensão dos fenômenos.

Assim, os dados não “falam por si só”. Dessa forma, toda pesquisa quantitativa seguiria o seguinte processo: qualitativo → quantitativo → qualitativo. No início da pesquisa, devemos usar teoria, construir hipóteses e planejar o estudo. No segundo momento, podemos utilizar um método quantitativo coletando dados por meio de algum instrumento (que se foi bem elaborado, também usou teoria). No final, temos que recorrer novamente à teoria para atribuir sentido e interpretar os dados. Obviamente existem pesquisas quantitativas que merecem ser criticadas por carecerem de uma base mais sólida. Algumas pesquisas educacionais de origem estadunidense, por exemplo, são bastante criticadas por possuírem técnicas estatísticas muito avançadas, mas serem pobres do ponto de vista teórico-interpretativo. No entanto, no atual nível de conhecimento científico, algumas críticas com relação aos métodos quantitativos do tipo “os estudantes não podem ser resumidos a números e tabelas” são consideradas no mínimo ingênuas. Além disso, determinadas críticas direcionadas às pesquisas quantitativas (ex: pobreza teórica) também são válidas para pesquisas qualitativas de baixa qualidade. Ou seja, não é o método que garante um trabalho bem realizado.

Finalmente, com relação à atual configuração do uso de métodos quantitativos na educação, Gatti (2004) realizou um criterioso levantamento sobre a produção brasileira nos últimos anos. Foram analisados todos os números publicados no período de 1970-2004 dos periódicos: Revista Brasileira de Estudos Pedagógicos (Inep/MEC), Cadernos de Pesquisa (Fundação Carlos Chagas), Educação e Realidade (UFRS), Educação e Sociedade (Cedes/Unicamp), Revista Brasileira de Educação (ANPEd), Estudos em Avaliação Educacional (Fundação Carlos Chagas), Ensaio (Fundação Cesgranrio). A autora chegou às seguintes conclusões: atualmente, na pesquisa educacional, poucos estudos empregam metodologias quantitativas; o uso de bases de dados existentes sobre educação é muito pequeno (isso seria uma evidência da dificuldade dos educadores em lidar com esse tipo de dados); existe uma dificuldade de leitura crítica dos trabalhos que utilizam métodos quantitativos; o uso de dados

quantitativos na pesquisa educacional no Brasil nunca teve uma tradição sólida; entre os poucos estudos que utilizavam métodos quantitativos, a maioria empregava apenas análise descritiva de tabelas de frequências, alguns poucos correlações e raríssimos estudos empregavam análise multidimensional (chegamos a essa mesma conclusão na nossa revisão da literatura sobre as concepções de avaliação dos alunos); os estudos quantitativos na área de educação que usam técnicas de análise mais sofisticadas na maioria das vezes não são realizados por educadores, e sim por profissionais com formação em outras áreas (isso pode ser considerado um indício de que a formação acadêmica na área da educação não tem contemplado os métodos quantitativos); no campo de estudos da avaliação educacional se encontra a maioria dos trabalhos quantitativos nos últimos anos, sendo nessa área que modelos de análise mais sofisticados vêm sendo usados (teoria da resposta ao item, modelos de análise hierárquica, estudos de relações multivariadas, entre outros).

Nesse sentido, acreditamos que nosso trabalho justifica-se pela necessidade de suprir uma lacuna existente nas pesquisas educacionais brasileiras, inclusive nas pesquisas sobre as concepções de avaliação dos alunos: o pouco uso de metodologias quantitativas, principalmente de métodos de análise de dados mais robustos. Por isso, decidimos analisar as concepções de avaliação dos alunos do ensino superior por meio de um questionário denominado *Students' Conceptions of Assessment - version VI - SCoA-VI* (BROWN, 2006) (ANEXO A), usando os seguintes métodos quantitativos: Escalonamento Multidimensional, Análise Fatorial Exploratória e Confirmatória, Análise de Variância, Modelagem de Equações Estruturais. Tais métodos se mostram adequados para retratar grandes coletivos, para conhecer grandes grupos de alunos ou escolas. O próximo capítulo descreve de forma mais detalhada o SCoA e o capítulo 5 apresenta todos os métodos quantitativos citados.

CAPÍTULO 3 - CARACTERIZAÇÃO E ADAPTAÇÃO DO *STUDENTS' CONCEPTIONS OF ASSESSMENT (SCOA)*

3.1 - Uma descrição do SCoA

O SCoA foi desenvolvido na Nova Zelândia e avalia 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 fortemente¹¹. O instrumento original em inglês possui 33 itens distribuídos em oito escalas: Melhora (aluno), Melhora (professor), Emoção (aluno), Emoção (sala), Irrelevante (ruim), Irrelevante (ignorar), Responsabilização (aluno), Responsabilização (escola)¹². As escalas do SCoA-VI avaliam as concepções de avaliação presentes na literatura educacional (ver seção 2.1), tornando cada uma delas mais específica por meio de subdivisões. Assim, o questionário divide a concepção de melhora em duas: o uso da avaliação por parte do aluno ou do professor para a melhora do ensino-aprendizagem. Os itens que investigam essa concepção são:

Melhora (aluno)

- 1- Eu presto atenção nos meus resultados de avaliação para me concentrar no que eu posso melhorar da próxima vez
- 10- Eu faço uso do *feedback* que recebo para melhorar 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
- 19- Eu uso as avaliações para identificar o que eu preciso estudar em seguida

Melhora (professor)

- 5- A avaliação ajuda os professores a acompanhar o meu progresso
- 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
- 23- Os meus professores usam a avaliação para me ajudar a melhorar

¹¹Os termos em inglês: *Strongly Disagree, Mostly Disagree, Slightly Agree, Moderately Agree, Mostly Agree, Strongly Agree*, respectivamente.

¹²Os nomes das escalas em inglês: *Improvement (self), Improvement (teacher), Affect (self), Affect (class), Irrelevant (bad), Irrelevant (ignore), Accountability (self), Accountability (school)*, respectivamente.

27- Os professores usam os meus resultados da avaliação para ver o que eles precisam me ensinar em seguida

30- A avaliação mostra se eu posso analisar e pensar criticamente sobre um assunto

O SCoA divide a concepção de que a avaliação tem um impacto emocionalmente positivo nos alunos em duas: o impacto no aluno individualmente e na sala de aula como um todo. Os itens que investigam essa concepção são:

Emoção (aluno)

6- A avaliação é uma experiência envolvente e agradável para mim

31- Eu realmente aprecio o aprendizado quando eu sou avaliado

Emoção (sala)

2- A avaliação encoraja a minha turma a trabalhar junta e a ajudar uns aos outros

12- A avaliação me motiva e aos meus colegas a ajudarem uns aos outros

17- A nossa turma se dá mais apoio quando nós somos avaliados

21- Quando nós fazemos avaliações, existe um bom clima na nossa turma

25- A avaliação faz a nossa turma cooperar mais uns com os outros

28- Quando nós somos avaliados, a nossa turma se torna mais motivada para aprender

O questionário divide a concepção de irrelevância em duas: considerar a avaliação como algo a ser ignorado ou como algo ruim. Os itens que investigam essa concepção são:

Irrelevante (ruim)

3- A avaliação é injusta com os alunos

13- A avaliação interfere no meu aprendizado

18- Os professores avaliam exageradamente

22- Os resultados da avaliação não são muito exatos

26- A avaliação é sem valor

Irrelevante (Ignorar)

7- Eu ignoro as informações da avaliação

29- Eu desconsidero os meus resultados de avaliação

32- A avaliação tem um impacto pequeno no meu aprendizado

O SCoA divide a concepção de responsabilização em duas: a avaliação torna os alunos ou as escolas responsáveis. Os itens que investigam essa concepção são:

Responsabilização (aluno)

- 4- Os resultados da avaliação mostram o quanto eu sou inteligente
- 16- Os resultados da avaliação predizem o meu desempenho futuro
- 20- A avaliação é importante para a minha carreira futura ou emprego

Responsabilização (escola)

- 11- A avaliação fornece informação sobre como as escolas estão indo
- 24- A avaliação mede a qualidade das escolas

Destacamos ainda que o SCoA, além de medir concepções de avaliação, também 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), uma lista de práticas avaliativas pode ser mostrada aos sujeitos. O SCoA faz isso por meio da “Parte Complementar do Questionário” (ANEXO A). A pergunta dessa parte é: “Quando você pensa na palavra AVALIAÇÃO, quais tipos de atividades avaliativas vêm à sua mente?”. Logo em seguida, são apresentadas diversas atividades avaliativas. Os alunos podem marcar quantas opções quiserem. São alguns exemplos: trabalho em grupo, autoavaliação, prova com questões abertas, dentre outros. Portanto, os alunos indicam suas definições de avaliação apontando as práticas avaliativas que eles associam com o termo avaliação.

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 (por isso um dos nossos objetivos específicos é analisar como as concepções de avaliação predizem as definições de avaliação dos alunos). Diversos estudos já utilizaram essa abordagem com professores (BROWN, 2002) e estudantes do ensino médio (BROWN; IRVING; PETERSON; HIRSCHFELD, 2009).

3.2 - Referencial teórico do SCoA

O desenvolvimento do *Students' Conceptions of Assessment - version VI* (SCoA-VI) (BROWN, 2006) foi baseado na teoria da aprendizagem auto-regulada (*self-regulated learning*). Esse também foi o referencial teórico utilizado no presente trabalho.

3.2.1 - Aprendizagem auto-regulada

A questão da auto-regulação na Psicologia aparece em várias abordagens teóricas. A análise do comportamento considera como comportamentos de auto-regulação estabelecimento de metas, auto-instrução, auto-monitoramento e auto-reforçamento. Considerando que um pressuposto básico dessa teoria é apontar que a consequência que se segue a um comportamento influencia a probabilidade de esse comportamento ocorrer novamente, a auto-regulação está relacionada com o uso e controle de estratégias comportamentais que levarão à obtenção de determinadas recompensas (SKINNER, 1993). Piaget destaca que os processos de auto-regulação da mente possibilitam que o sujeito se adapte ao meio. Assim, pode-se chamar adaptação ao equilíbrio dos processos de assimilação e acomodação. Na teoria Piagetiana, toda ação, exterior ou mesmo inteiramente interior, corresponde a uma necessidade. Toda necessidade tende a incorporar as coisas e pessoas à atividade própria do sujeito, isto é, assimilar o mundo exterior às estruturas já construídas. Essas estruturas, por sua vez, são reajustadas em função das transformações ocorridas, ou seja, acomodadas aos objetos externos. Assimilando assim os objetos, a ação e o pensamento são compelidos a se acomodarem a estes (PIAGET, 1990). Vygotsky enfatiza a construção sócio-histórica do desenvolvimento, apontando que a auto-regulação e o desenvolvimento das funções psíquicas superiores têm origem social. Para ele, é durante o início da idade escolar que as funções intelectuais superiores, cujas características são a consciência reflexiva e o controle deliberado, adquirem um papel fundamental no desenvolvimento. A auto-regulação, portanto, está relacionada com a tomada de consciência dessas funções. Consciência e controle são muito importantes para a aprendizagem. Isso explica em parte a dificuldade de auto-regulação de crianças muito novas (VYGOTSKY, 1990; 1993).

Embora existam diferentes definições de aprendizagem auto-regulada em função das orientações teóricas dos pesquisadores, uma conceituação comum pode ser apontada: o uso sistemático de estratégias metacognitivas, motivacionais e comportamentais é uma característica chave da maioria das definições de alunos auto-regulados. Em relação aos

processos metacognitivos, alunos auto-regulados planejam, estabelecem objetivos, organizam, auto-monitoram, e auto-avaliam em vários pontos durante o processo de aprendizagem. Considerando os processos motivacionais, os alunos auto-regulados possuem crenças altas de auto-eficácia e interesse intrínseco à tarefa. Para os observadores, eles são alunos auto-iniciados que demonstram um esforço extraordinário e persistência durante o aprendizado. Já em relação aos processos comportamentais, alunos auto-regulados selecionam, estruturam e criam ambientes que otimizam o aprendizado. Eles se auto-instruem e se auto-reforçam, procuram avisos, informações e lugares onde estão mais aptos a aprender (ZIMMERMAN, 1990).

O processo de auto-regulação está presente em todas as pessoas ao longo de toda a vida. Esse processo opera por meio de três subfunções psicológicas: a) auto-observação: fornece as informações necessárias que usamos como guias, como referências nos nossos processos de desempenho e julgamento sobre o comportamento em curso; b) processos de julgamento: servem para extrair padrões genéricos das várias avaliações de diferentes atividades e ambientes, processar as informações e formar um novo padrão geral que será usado pelo próprio sujeito; c) auto-reação: serve para motivar e regular o curso das ações, para possibilitar a mudança autodirigida do comportamento, seja por consequências reforçadoras, punitivas ou indiferentes. Apesar desse ciclo, o processo de auto-regulação não é mecânico (POLYDORO; AZZI, 2008).

Nessa perspectiva, o homem é considerado como sendo capaz de exercer um papel ativo em sua própria história (concepção elaborada sob o nome de agência pessoal). Isso ocorre devido a quatro principais capacidades humanas: a) simbolização: representação e interpretação das próprias experiências e desempenhos, atribuindo significado individual aos mesmos; b) pensamento antecipatório: transcender o presente imediato e regular o comportamento atual em direção a objetivos e metas futuras; c) auto-regulação: monitoramento, avaliação e controle pelo indivíduo de seu próprio comportamento, em direção aos objetivos pessoais; d) auto-reflexão: analisar experiências e processos de pensamentos, bem como organizar auto-percepções. Além disso, o pensamento e ação humana são considerados produtos de uma inter-relação dinâmica entre influências pessoais, comportamentais e ambientais (BANDURA, 2001).

Assim, a teoria da aprendizagem auto-regulada considera que a auto-regulação do aprendizado envolve escolha, tomada de decisão, planejamento e responsabilidade por nossas ações (BORUCHOVITCH, 2004). Alunos auto-regulados são estudantes ativos que gerenciam de uma forma eficaz o seu próprio processo de aprendizagem, assim como a sua motivação. Estabelecem metas para eles mesmos, monitorando suas

estratégias de aprendizagem e sua motivação em função das exigências específicas de cada tarefa (BZUNECK, 2001).

Nesse sentido, um aspecto importante das teorias de aprendizagem auto-regulada é que o aprendizado do estudante e a motivação são tratados como processos interdependentes. Não podemos compreender um processo sem fazer referência ao outro. Para Bzuneck (2001), as crenças de auto-eficácia estão entre os fatores que integram os mecanismos psicológicos da motivação do aluno. Elas podem ser definidas como um julgamento das próprias capacidades de realizar ações exigidas para se atingir um determinado grau de desempenho (BANDURA, 1986). Não se trata de alguém possuir ou não certas capacidades. O que é colocado em questão quando se analisam as crenças de auto-eficácia é o fato de a pessoa acreditar que possua tais capacidades. Trata-se de uma avaliação, de uma percepção pessoal (BZUNECK, 2001). Assim, essas crenças de competência pessoal fornecem a base para a motivação humana e as realizações pessoais. Ou seja, a menos que as pessoas acreditem que suas ações possam produzir os resultados desejados, elas terão pouca motivação para agir ou perseverar diante de dificuldades. As crenças e a realidade nunca possuem um encaixe perfeito, sendo que as pessoas geralmente são orientadas por suas crenças quando se relacionam com o mundo. Além disso, como os indivíduos agem de forma individual e coletiva, a auto-eficácia pode ser entendida como um construto pessoal e social. Dessa forma, sistemas coletivos também desenvolvem um sentido de eficácia coletiva, de uma crença compartilhada pelo grupo em sua capacidade de atingir objetivos e realizar tarefas. As escolas, por exemplo, desenvolvem crenças coletivas sobre a capacidade dos alunos de aprender, dos professores de ensinar e dos gestores de criar condições para o processo de ensino-aprendizagem (PAJARES; OLAZ, 2008). Os colegas de uma sala de aula também desenvolvem crenças coletivas.

Em síntese, a teoria da aprendizagem auto-regulada destaca a importância da iniciativa pessoal na aprendizagem. Ao contrário de alunos passivos, alunos auto-regulados procuram de uma maneira pró-ativa informações quando precisam e assumem os passos necessários para isso. Quando encontram obstáculos como condições inadequadas de estudo, livros ruins ou professores confusos eles desenvolvem estratégias para aprender. Além disso, alunos auto-regulados assumem grande responsabilidade pelos seus resultados. A aprendizagem auto-regulada, portanto, analisa como os estudantes selecionam, organizam, ou criam ambientes de aprendizagem vantajosos para eles mesmos. Sem dúvida, todos os estudantes usam processos regulatórios em algum grau, mas alunos auto-regulados são distintos por: sua

consciência da relação estratégica entre processos regulatórios ou respostas e resultados de aprendizagem; seu uso dessas estratégias para conseguir seus objetivos acadêmicos (ZIMMERMAN, 1990).

Os resultados do questionário SCoA, portanto, são interpretados a partir da teoria da aprendizagem auto-regulada. Estudantes que assumem responsabilidade por suas próprias ações geralmente atingem melhores resultados na maioria das medidas educacionais. Já aqueles que não assumem responsabilidade por suas ações tendem a atingir piores resultados educacionais. 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.

Ainda com relação à ligação entre as concepções de avaliação e a teoria da aprendizagem auto-regulada, Brown, Peterson e Irving (2009) discutem a hipótese de que somente a concepção de melhora parece estar alinhada com uma resposta positiva de auto-regulação 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 auto-regulação do aprendizado.

Por fim, existem várias modalidades de avaliação que podem ser classificadas pelo momento em que acontecem e pelo objetivo que possuem. São elas: avaliação diagnóstica inicial (antes do processo ensino-aprendizagem); avaliação formativa (durante o processo ensino-aprendizagem); avaliação somativa (depois do processo ensino-aprendizagem) (SANT’ANNA, 1995). Essa última é a forma de avaliação mais conhecida dos profissionais da educação, que geralmente associam a avaliação com os resultados alcançados no final do processo de ensino. Já com relação à avaliação formativa, sintetizamos sua definição e objetivos no trecho a seguir:

A avaliação durante o processo de aprendizagem ou avaliação formativa é um termo introduzido em 1967 por M. Scriven para se referir aos procedimentos utilizados pelos professores para adaptar

seu processo didático aos progressos e necessidades de aprendizagem observados em seus alunos. [...] Pode-se dizer, pois, que a avaliação formativa destaca a regulação das atuações pedagógicas e, portanto, interessa-se fundamentalmente mais pelos procedimentos das tarefas do que pelos resultados. Em resumo, a avaliação formativa persegue os seguintes objetivos: a regulação pedagógica, a gestão dos erros e a consolidação dos êxitos (JORBA; SANMARTÍ, 2003, p. 30).

Alguns itens do SCoA levam em consideração essa função formativa, de regulação das atividades pedagógicas da avaliação. São exemplos os itens: “Os professores usam os meus resultados da avaliação para ver o que eles precisam me ensinar em seguida” e “Eu observo o que eu fiz de errado ou de maneira insuficiente para guiar o que eu deveria aprender em seguida”.

3.3 - O contexto e resultados de pesquisas desenvolvidas na Nova Zelândia

Existe uma concordância de que a avaliação tem pelo menos três propósitos principais: melhorar o ensino e o aprendizado; tornar escolas, professores e alunos responsáveis pelo aprendizado; e uma rejeição de todos esses propósitos (a avaliação é irrelevante). Pesquisas neozelandesas com alunos do ensino médio (BROWN; HIRSCHFELD, 2008; BROWN *et al.*, 2009; BROWN; IRVING; PETERSON, 2008) sobre concepções de avaliação identificaram propósitos adicionais: a avaliação tem um impacto afetivo ou um benefício emocional. Assim, essa concepção de que a avaliação tem um impacto emocionalmente positivo nos alunos foi dividida em duas escalas relacionadas: o impacto emocional no aluno individualmente, o prazer pessoal (Emoção-aluno) e o benefício emocional da sala de aula como um todo (Emoção-sala).

Pesquisas desenvolvidas na Nova Zelândia mostraram que: 1) os alunos têm baixos níveis de prazer pessoal na avaliação (Emoção-aluno); 2) “Emoção-aluno” e “Emoção-sala” estão associadas a um desempenho acadêmico mais baixo em leitura e matemática (BROWN; HIRSCHFELD, 2007, 2008) e têm relações estatisticamente invariantes com leitura através da idade, sexo e etnia dos estudantes (HIRSCHFELD; BROWN, 2009). Embora pudéssemos esperar, de acordo com uma perspectiva do senso comum, que um impacto emocionalmente positivo nos alunos (individualmente ou na sala de aula) poderia elevar o desempenho acadêmico, parece que essa concepção de avaliação é desadaptativa¹³. Brown e Hirschfeld (2008) sugeriram que os estudantes que

¹³ O termo em inglês é *maladaptive*. Encontramos como possíveis traduções para essa palavra: inadaptado, inadequado e desajustado. Optamos por traduzir como “desadaptativo”, mesma tradução utilizada por Cláudio (2009). Essa parece ser a tradução adotada na área de pesquisa psicológica.

priorizam tal concepção provavelmente não levam a avaliação a sério e assim não prestam atenção ou se esforçam o suficiente para se sair bem.

Brown *et al.* (2009) relataram que alunos neozelandeses do ensino médio identificaram dois grandes grupos de atividades avaliativas: práticas formais controladas pelo professor do tipo exames e testes; práticas interativo-informais como a autoavaliação. Os itens na categoria de avaliações formais controladas pelo professor foram selecionados três vezes mais do que aqueles na categoria interativo-informal. Isso indicou claramente que os estudantes associaram a avaliação com práticas formais do tipo exames e testes. Além disso, outros resultados também foram encontrados: a) práticas formais controladas pelo professor foram preditas negativamente por “Emoção (aluno)” (ou seja, exames não eram agradáveis); b) práticas interativo-informais foram preditas positivamente por “Emoção (sala)” e “Irrelevante (ignorar)”; c) práticas formais controladas pelo professor foram preditas positivamente por “Melhora (professor)”. Dessa forma, os autores concluíram que os alunos pareciam pensar: *“Se a avaliação é controlada pelo professor, eu não necessariamente vou gostar disso, mas isso vai ajudá-lo a me ensinar melhor. As avaliações interativas são boas para a dinâmica da sala de aula, mas eu não preciso prestar muita atenção nelas.”*

Brown, Irving, e Peterson (2008) encontraram o mesmo padrão de relações entre o prazer pessoal e o benefício emocional da sala de aula com as definições de avaliação formal e informal. As práticas interativo-informais não apresentaram uma associação estatisticamente significativa com o desempenho acadêmico, enquanto as práticas formais foram capazes de prever um aumento no desempenho acadêmico. Assim, parece que entre os estudantes neozelandeses o efeito de enfatizar um impacto emocionalmente positivo nos alunos e na sala é fundamentalmente desadaptativo com relação ao desempenho acadêmico.

Em síntese, podemos dizer que as pesquisas desenvolvidas na Nova Zelândia têm como foco diferenças individuais e desempenho dos alunos (relaciona as concepções de avaliação com a performance acadêmica dos alunos), além de possuir como público alvo principal estudantes do ensino médio. No caso da nossa pesquisa, trabalhamos com alunos do ensino superior e não tivemos como foco o desempenho acadêmico dos alunos (consideramos que antes de estabelecer relações entre concepções de avaliação e desempenho acadêmico, seria mais prudente realizar uma primeira pesquisa que investigasse quais concepções seriam encontradas entre os estudantes brasileiros).

Essas considerações nos remetem para possíveis diferenças culturais entre o Brasil e a Nova Zelândia. Dessa forma, destacamos que a possibilidade de modelos alternativos para a realidade brasileira foi investigada na presente pesquisa.

3.4 - Diferenças culturais, tradução e validação de instrumentos

Realizar a tradução e a adaptação de questionários para línguas e culturas diferentes não é uma tarefa simples, pois várias etapas precisam ser desenvolvidas¹⁴. Precisamos tomar alguns cuidados para não fazer uma aplicação ingênua ou acrítica de conceitos e estruturas provenientes de uma determinada cultura para outra (MATOS, 2006). Devemos levar em conta as características socioeconômicas e culturais das amostras em que os instrumentos serão aplicados. A utilização de um instrumento estrangeiro sem uma adaptação criteriosa pode colocar em risco a validade e precisão de avaliações realizadas (NASCIMENTO; FIGUEIREDO, 2002). Diversos autores destacam essa discussão. Aldridge *et al.* (1999), por exemplo, apontam a necessidade de cuidado na interpretação de resultados de instrumentos aplicados em países com diferenças culturais. Walker (2004) evidencia a influência de fatores socioculturais e recomenda cautela no uso de instrumentos geralmente oriundos do Ocidente em investigações transculturais. No entanto, apesar das dificuldades apontadas, os estudos transculturais vêm se apresentando como uma possibilidade de pesquisa rica e promissora na educação e na psicologia (MATOS, 2006).

Um dos argumentos mais importantes para utilizar o mesmo instrumento em países diferentes é a questão da comparação. Somente assim se torna possível a análise dos resultados com uma interpretação comum. Isso assegura uma equivalência das escalas e dos itens e agrega dados para a validação do instrumento (MATOS, 2006).

Com o intuito de auxiliar os pesquisadores no processo de adaptação de testes para diversas línguas e culturas, um comitê internacional denominado Comissão Internacional de Testes (*International Test Commission - ITC*) criou várias diretrizes para a adaptação e uso de testes psicológicos e educacionais em contextos linguísticos e culturais diferentes (MATOS, 2006). Partindo das recomendações da ITC, pesquisadores como Hambleton e Patsula (1999) sugerem vários passos para adaptar um teste de uma cultura para outra, tais como: decidir entre a adaptação de um teste existente ou o desenvolvimento de um novo; selecionar tradutores qualificados;

¹⁴ Para uma discussão detalhada sobre linguagem, tradução, diferenças culturais, adaptação e validação de instrumentos, ver Matos (2006).

documentar o processo de adaptação e preparar um manual para os usuários do teste; realizar um treinamento com os usuários; entre outros.

Assim, considerando a complexidade envolvida no processo de tradução e adaptação de questionários para línguas e culturas diferentes, apresentamos a seguir um referencial teórico para o processo de adaptação e validação do SCoA para a realidade brasileira.

Primeiramente, no processo de tradução de questionários, existe uma técnica conhecida como tradução inversa (*back translation*). Esse método se refere à tradução de uma tradução de volta a língua de origem. A tradução inversa geralmente é utilizada para avaliar a qualidade da tradução, já que ela compara e contrasta a tradução inversa com o texto original. Dessa forma, a tradução inversa é uma parte do processo de tradução do questionário (HARKNESS; SCHOUA-GLUSBERG, 1998).

Existem vários modelos disponíveis para a validação de instrumentos. Utilizamos na presente pesquisa uma proposta que se baseia em três grandes polos: procedimentos teóricos, procedimentos empíricos (experimentais) e procedimentos analíticos (estatísticos) (PASQUALI, 1999).

Os procedimentos teóricos se iniciam com a análise teórica dos itens. Nessa etapa inicial, estamos diante da hipótese de que os itens representam de maneira adequada os construtos do instrumento. Nessa fase isso é avaliado por meio da opinião de outros em dois procedimentos: análise dos juízes e análise semântica. A análise dos juízes consiste em solicitar que especialistas da área avaliem os itens do questionário e sua adequação aos respectivos construtos (nessa pesquisa, não utilizamos essa técnica). Já a análise semântica dos itens é um método que tem como objetivo verificar se todos os itens são compreensíveis para os membros da população alvo (PASQUALI, 1999). Uma das maneiras mais eficazes de testar a compreensão dos itens é verificá-la em uma situação de entrevista com pequenos grupos, com três a quatro pessoas (PASQUALI, 2003).

Os procedimentos empíricos (experimentais) de validação do instrumento envolvem o planejamento da aplicação e a coleta da informação empírica. O planejamento da aplicação envolve a definição da amostra (tamanho e público alvo) e das instruções de aplicação. Já na coleta da informação, é necessário seguir exigências presentes na aplicação de qualquer instrumento psicológico, como aplicar o teste em um ambiente tranquilo e sem distrações e treinar os aplicadores (PASQUALI, 1999).

Por fim, os procedimentos analíticos (estatísticos) são complementares às diversas etapas anteriores de validação do questionário. Os procedimentos de tradução e adaptação têm que ser acompanhados de investigações estatísticas das propriedades de

medida do instrumento (HARKNESS; SCHOUA-GLUSBERG, 1998). Além disso, no desenvolvimento e validação de instrumentos psicométricos de acordo com critérios científicos, precisamos trabalhar com amostras grandes. Métodos estatísticos multivariados geralmente fazem exigências importantes dos dados, que devem produzir uma variância suficiente para que a análise seja consistente. No caso da análise fatorial, por exemplo, qualquer análise com menos de 200 sujeitos dificilmente pode ser levada em consideração (PASQUALI, 1999).

Nesse sentido, utilizamos todos esses procedimentos na adaptação e validação do SCoA. Esse instrumento foi usado para investigar diversas questões sobre o ensino superior, como, por exemplo, as concepções de avaliação dos alunos e as diferenças entre uma IES pública e uma IES privada. O próximo capítulo descreve de forma mais detalhada todos os objetivos e hipóteses da pesquisa.

CAPÍTULO 4 - OBJETIVOS E HIPÓTESES DA PESQUISA

4.1 - Objetivos

4.1.1 - Objetivo Geral

- Analisar as concepções de avaliação dos alunos do Ensino Superior.

4.1.2 - Objetivos específicos

- Adaptar e validar o questionário *Students' Conceptions of Assessment* (SCoA) - *version VI* - para a realidade brasileira.
- Investigar as definições de avaliação dos alunos do Ensino Superior.
- Analisar como as concepções de avaliação predizem as definições de avaliação dos alunos.
- Identificar e analisar possíveis diferenças nas concepções de avaliação dos alunos pertencentes a uma IES pública e a uma IES privada (uma Universidade Federal e um Centro Universitário privado, ambas IES do estado de Minas Gerais).

4.2 - Hipóteses

Considerando tudo o que foi apontado e discutido nos três primeiros capítulos, estabelecemos as seguintes hipóteses para essa pesquisa:

- Hipótese 1 - concepções de avaliação (H1): Existem múltiplas concepções de avaliação dos alunos do ensino superior brasileiro (tendo como base principalmente a revisão da literatura sobre as concepções de avaliação dos alunos universitários brasileiros e a literatura e pesquisas internacionais sobre as concepções de avaliação dos alunos).
- Hipótese 2 - definições de avaliação (H2): Os alunos do ensino superior identificam dois grandes grupos de atividades avaliativas: práticas formais e práticas interativo-informais (tendo como base principalmente as pesquisas de Brown *et al.*, 2009; Brown, Irving, e Peterson, 2008).
- Hipótese 3 - definições de avaliação (H3): Os alunos do ensino superior definem a avaliação principalmente como métodos formais/convencionais do tipo exames e testes (tendo como base a revisão da literatura brasileira e as pesquisas de Brown *et al.*, 2009; Brown, Irving, e Peterson, 2008).

- Hipótese 4 - concepções de avaliação predizendo as definições de avaliação dos alunos do ensino superior (H4): As práticas formais são preditas negativamente por “Emoção (aluno)”; as práticas interativo-informais são preditas positivamente por “Emoção (sala)” e “Irrelevante (ignorar)”; as práticas formais são preditas positivamente por “Melhora (professor)”. (tendo como base principalmente a pesquisa de Brown *et al.*, 2009).
- Hipótese 5 - diferenças entre IES pública e IES privada (H5): Não existem diferenças nas concepções de avaliação dos alunos quando comparamos instituições de ensino públicas e privadas (não encontramos elementos na nossa revisão da literatura que nos indiquem possíveis diferenças). No entanto, destacamos que a nossa experiência como professor de IES privadas e como professor e estudante de IES públicas nos fez pensar inicialmente em possíveis diferenças.

CAPÍTULO 5 - MÉTODO

5.1 - O desenho de pesquisa

Essa pesquisa teve um delineamento de *survey* (*survey design*). 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. O delineamento de *survey* difere de uma pesquisa experimental (*experimental design*) porque não envolve um tratamento dado aos participantes pelo pesquisador (CRESWELL, 2005). Entretanto, no caso de investigações não experimentais, uma série de métodos estatísticos pode ser usada para responder questões causais e não-causais. Técnicas que respondem questões não-causais incluem Análise Fatorial Exploratória e Confirmatória, enquanto Modelagem de Equações Estruturais e Análise Regressão tentam responder questões causais. Quando comparada com um delineamento correlacional (*correlational design*), a pesquisa de *survey* tem muito em comum, pois pesquisadores que usam *survey* frequentemente correlacionam variáveis.

Embora muitas formas diferentes de *survey* existam, os pesquisadores geralmente coletam dados usando duas formas básicas: questionários e entrevistas. A distinção entre eles pode ser feita examinando quem completa os dados no instrumento: os participantes ou o pesquisador. No questionário, o participante de um estudo completa e retorna o instrumento ao pesquisador. Já na entrevista, o pesquisador anota as respostas fornecidas pelo participante do estudo. Esse procedimento de entrevista quantitativa não deve ser confundido com uma entrevista qualitativa (CRESWELL, 2005).

No presente trabalho, fizemos uso de um questionário para a coleta de dados. Nesse ponto, uma questão da terminologia precisa ser discutida. A literatura nacional e internacional de educação frequentemente trata como sinônimos os termos *survey*, questionário e escala. Já fizemos uma primeira distinção ao apontar que o questionário é uma das formas de *survey*. Entretanto, discutimos a seguir uma distinção pouco conhecida que se refere à diferença entre Escala Psicométrica e *Survey*. Segundo Gall, Gall e Borg (2007) existem as seguintes diferenças:

Escala Psicométrica:

- Vários itens para medir cada construto latente (não observável diretamente).
- Os itens que medem o mesmo construto são altamente correlacionados.

Survey:

- Os itens não estão necessariamente relacionados uns com os outros.
- Os itens podem medir características diretamente observáveis.
- Atitudes são medidas por um único item ou por poucos itens.

Acreditamos que tratar os termos *survey*, questionário e escala como sinônimos reflete uma dificuldade teórica de diferenciar instrumentos de pesquisa. Assim, para sermos conceitualmente mais precisos, destacamos que o *Students' Conceptions of Assessment* (SCoA) - *version VI* é uma Escala Psicométrica. Considerando uma nomenclatura da *American Psychological Association* (APA), o SCoA seria denominado um teste. Essa inclusive é a maneira como o autor do SCoA cita esse instrumento: “Brown, G. T. L. (2003). *Students' conceptions of assessment (SCoA) inventory* (Versions 1-6). Unpublished test. Auckland, NZ: University of Auckland”. No entanto, optamos por usar o termo questionário nessa pesquisa por ser um termo que aparece com mais frequência na literatura e, conseqüentemente, é mais inteligível para a maioria das pessoas.

Vale lembrar também a definição que demos nessa pesquisa para o termo concepção na seção 2.1.1: “sistemas complexos de explicação”, uma “rede complexa de ideias, conceitos, representações e preconceitos”. De forma complementar, também entendida como: “operação de construção de conceitos”, que são uma “representação, ideia, geral e abstrata”; informação sobre a maneira como as pessoas percebem, avaliam e agem com relação ao fenômeno da avaliação. Considerando que o pesquisador precisa estar consciente do que está fazendo e quais são suas razões, acreditamos que os aspectos presentes nessas definições só podem ser medidos adequadamente com estímulos múltiplos (itens). Por isso usamos a abordagem psicométrica. Alguns fenômenos podem ser medidos por um único item como, por exemplo, o tipo de escola (pública ou privada). No entanto, no tipo de pesquisa psicológica que fazemos precisamos de indicadores múltiplos. Isso evidencia a natureza complexa do problema que estudamos: o pensamento dos alunos.

Nesse sentido, outra questão do nosso delineamento de pesquisa está relacionada com a amostra e suas características. Na educação, muitas vezes o pesquisador não pode

designar aleatoriamente participantes para grupos de um experimento (desenho experimental). Isso pode acontecer por causa da disponibilidade dos sujeitos ou porque a situação não permite que sejam formados grupos artificiais. Os educadores frequentemente usam grupos intactos como escolas e salas de aula (CRESWELL, 2005). Destacamos também outras dificuldades associadas ao trabalho com amostras aleatórias como o fator financeiro, acesso aos sujeitos e tempo para a realização da pesquisa. Por exemplo: não tínhamos acesso irrestrito a todos os alunos das duas IES que investigamos. Contatamos muitos cursos, mas só tivemos acesso liberado a alguns deles. Assim, a amostra da presente pesquisa não foi aleatória. Foi uma amostragem não-probabilística. A principal consequência disso é a impossibilidade de generalização dos resultados.

Ainda com relação ao desenho de pesquisa, considerando que não podemos monitorar diretamente o pensamento humano, precisamos utilizar procedimentos inferenciais como o auto-relato (*self-report*) e a observação. Auto-relatos são geralmente obtidos por meio de entrevistas, discussões em pequenos grupos ou questionários (BROWN, 2008). Diversos tipos diferentes de questionários e entrevistas são usados na pesquisa quantitativa de *survey*. Os tipos mais usados na educação são: questionários enviados por correio, questionários eletrônicos, entrevista face-a-face, entrevista de grupo focal e entrevista por telefone¹⁵ (CRESWELL, 2005).

Assim como todos os outros métodos, os métodos de *survey* apresentam vantagens e limitações. Questionários são inflexíveis e usualmente requerem um nível razoável de letramento e compreensão de linguagem instrumental (FOWLER, 1993; HUGHES, 1980). Outras limitações do uso de questionários: índice baixo de retorno; presença de respostas em branco ou incompletas; impossibilidade de resolução de dúvidas pela ausência do pesquisador (uma das maneiras de amenizar os aspectos negativos é enviar junto com o questionário um texto explicativo, onde podem ser colocados: instruções para responder ao questionário, objetivos da pesquisa, questões relacionadas a sigilo, prazo para envio das respostas, importância da participação, forma de utilização das respostas) (MACHADO; MAIA; LABEGALINI, 2007); esforço considerável para escrever questões que possam ser compreendidas da mesma forma por todos os participantes e para a validação do instrumento. Obviamente, algumas dessas críticas se referem a tipos diferentes de trabalho com questionário. Por exemplo: existem situações onde o pesquisador aplica o questionário presencialmente.

¹⁵ Os termos em inglês são: *mailed questionnaires*, *electronic questionnaires*, *one-on-one interviews*, *focus group interviews*, *telephone interviews*, respectivamente.

Por outro lado, também existem pontos favoráveis no uso de questionários: pode fornecer dados atualizados sobre a temática investigada; pode coletar um volume significativo de dados em um tempo curto; não exige pessoal treinado; atinge muitos sujeitos ao mesmo tempo; respeita o anonimato e por isso pode estimular cooperação; como o pesquisador está ausente na hora da resposta, isso pode diminuir influências e distorções dos dados fornecidos; custo operacional pequeno (MACHADO; MAIA; LABEGALINI, 2007). O questionário permite: abordar de uma forma padronizada todos os aspectos relevantes teoricamente e previamente identificados de um tópico de interesse (FOWLER, 1993); comparar resultados de uma maneira confiável e econômica; apresentar estímulos múltiplos para um mesmo conceito. Além disso, aplicações em larga escala reduzem a possibilidade de resultados ao acaso (BROWN, 2008). Novamente, algumas dessas vantagens se referem a tipos específicos de questionário (ex: não exigir pessoal treinado).

Quanto à aplicação do questionário, na nossa pesquisa fizemos uma aplicação em grupo auto-administrada (*group administration of self-administered surveys*) (DILLMAN, 2000). Ou seja: apesar do instrumento ser auto-administrado (o próprio aluno respondia), o questionário foi aplicado em grupo e presencialmente nas respectivas salas de aula dos estudantes. Mesmo com a grande economia de tempo e dinheiro que esse método propicia, aplicações em grupo podem exercer algumas influências que não acontecem em outros tipos de situações auto-administradas: comentários feitos pelo aplicador podem influenciar as respostas dos participantes; o estabelecimento de condições inapropriadas para responder ao instrumento (ex: reservar os últimos minutos de uma aula e dizer aos sujeitos que assim que terminarem o questionário podem ir embora. Isso pode gerar um comportamento de responder com pressa e de maneira desatenta); o contexto de uma sala de aula pode desencadear um comportamento nos alunos de resposta similar ao de uma prova (DILLMAN, 2000). Por causa de questões como essas, a literatura costuma indicar uma estruturação da aplicação em grupos fazendo uso, por exemplo, de um protocolo de aplicação com instruções. Considerando tais recomendações, elaboramos uma primeira página para o SCoA com instruções que eram lidas no momento da aplicação (ANEXO A). Nosso objetivo foi criar um ambiente de aplicação similar para todos os grupos. A versão original em inglês do SCoA não possui uma primeira página com instruções.

5.1.1 - Escalas de resposta

Uma questão problemática envolvida com o uso de questionários se relaciona com a maneira na qual os sujeitos devem responder. Alguns exemplos de formatos de resposta incluem concordância, frequência, importância, qualidade e probabilidade (GABLE; WOLF, 1993), sendo que a escolha da escala depende do tipo de fenômeno que está sendo investigado. Formatos de resposta que demandam que as pessoas relatem a frequência com que elas se envolvem em uma atividade estão propensos a muitos erros (BROWN, 2008). Algumas questões que afetam nossa memória e são possíveis fontes de erro nos auto-relatos dos sujeitos: nós esquecemos coisas no decorrer do tempo; nós não prestamos atenção a tudo que fazemos; às vezes não nos lembramos de algumas coisas; nossa memória pode ser mudada como consequência de influências externas; nossa memória pode sofrer influências inconscientes (SCHACTER, 1999); dentre outras. Ou seja: nossa memória é incompleta, seletiva e interpretativa, sendo que o intervalo de tempo envolvido (e eventos que aconteceram nesse intervalo) também pode influenciar. Além disso, as pessoas podem responder de forma diferente se a escala for em horas, dias, semanas ou meses. Assim, devemos estar atentos às nossas escolhas, pois formatos de resposta diferentes (escala *Likert*, escala diagramática com diferencial semântico, etc.) também possuem propriedades psicométricas diferentes, afetando inclusive a análise de dados (OFIR; REDDY; BECHTEL, 1987).

Existem evidências de que formatos de concordância estão menos sujeitos a erros de memória porque eliciam informações baseadas no presente, ao invés de exigir que os sujeitos recuperem conhecimentos, crenças ou atitudes (BROWN, 2008). Na presente pesquisa, a concordância foi o formato de resposta utilizado (o quanto um sujeito concorda ou discorda com uma afirmação ou concepção).

No entanto, mesmo após a decisão sobre o tipo de escala de resposta, outras decisões problemáticas se apresentam ao pesquisador com relação ao formato de resposta: o número de pontos de resposta, o nome (rótulo) dos pontos de resposta e a estrutura de pontos de resposta positivos e negativos. A literatura apresenta algumas considerações a esse respeito: um número de pontos de resposta entre 5 a 7 parece ser o ideal (COMREY; MONTAG, 1982); pontos neutros/médios¹⁶ devem ser evitados (ORY; WISE, 1981); as opções de resposta intermediárias devem ser nomeadas com o intuito de obter discriminações melhores (escalas nomeadas/rotuladas de forma pobre -

¹⁶ Exemplo: na escala *Likert* - quase nunca, raramente, às vezes, frequentemente, quase sempre - existe um ponto médio/neutro (às vezes). O argumento contra esse ponto neutro na literatura diz que ele pode funcionar com uma “fuga” para o sujeito que, na dúvida, acaba escolhendo essa opção.

ex: só a primeira e a última categorias são nomeadas - exigem mais esforço cognitivo na hora de responder) (LAM; KLOCKARS, 1982), dentre outras.

Geralmente, os questionários usam um formato de resposta balanceado no qual as categorias de resposta são simplesmente um espelho (exemplo: discordo fortemente e discordo são balanceados com concordo fortemente e concordo). Uma das fraquezas desse formato se refere ao fato de que, se os sujeitos estão inclinados a avaliar positivamente um determinado objeto, a escala restringe a variância das respostas de forma significativa. Isso está diretamente relacionado com o uso de técnicas estatísticas como a Análise Fatorial e a Modelagem de Equações Estruturais: precisamos de variação nas respostas para fazer um uso adequado de tais técnicas (BROWN, 2008). Na pesquisa social, estamos interessados nas relações entre as variáveis. Sem variação não vamos conseguir explicar as respostas, pois não somos capazes de explicar diferenças que não aparecem na coleta de dados. Assim, ao procurar relações entre variáveis, é necessário obter variância nas respostas. Para explicar a variância, devemos tentar maximizá-la entre os respondentes (BABBIE, 1999).

Nesse sentido, uma alternativa metodológica são os chamados formatos de resposta assimétricos (*skewed response formats*), que são escalas que têm mais opções de resposta positivas (*positively-packed*) ou que têm mais opções de resposta negativas (*negatively-packed*). No questionário do presente trabalho, foi utilizado um sistema de respostas balanceado positivamente, uma escala *Likert*¹⁷ de seis pontos: discordo fortemente, discordo na maior parte, concordo ligeiramente, concordo moderadamente, concordo na maior parte, concordo fortemente. Essa escala possui duas respostas negativas e quatro respostas positivas. Esse sistema de respostas balanceado positivamente tem se mostrado efetivo para gerar variância nas respostas quando os participantes são inclinados a ser orientados positivamente com relação a um construto (BROWN, 2004). O uso desse tipo de escala se justifica porque as pessoas parecem concordar com concepções múltiplas e contraditórias de avaliação ao mesmo tempo (ver seção 2.1). Assim, como existem evidências teóricas e empíricas para esperar que os alunos estivessem inclinados a concordar com todas as concepções, um sistema de respostas balanceado positivamente foi utilizado para ter uma maior variância das respostas ao SCoA.

¹⁷ O método de medida desenvolvido por Rensis Likert, denominado escala Likert, é um formato de resposta muito usado nos questionários de *survey*. O valor singular desse formato é a ordinalidade não ambígua das categorias de resposta (BABBIE, 1999).

5.2 - Métodos estatísticos

Passamos a explicar a seguir os métodos estatísticos que utilizamos em nossa pesquisa.

5.2.1 - Níveis de mensuração das variáveis

Antes de explicar os métodos estatísticos propriamente ditos, discutiremos alguns aspectos importantes sobre os tipos de variáveis. Existem dois tipos básicos de dados: não-métricos (qualitativos) e métricos (quantitativos). As medidas não-métricas podem ser efetuadas com uma escala nominal ou ordinal. A escala nominal designa números utilizados simplesmente para rotular indivíduos ou objetos. Assim, os números ou símbolos designados aos objetos não possuem significado quantitativo além de indicar a presença ou ausência de um determinado atributo. São exemplos: a religião ou partido político de um grupo de indivíduos. A escala ordinal reflete uma ordenação entre as categorias que constituem uma variável. Embora essas medidas sejam frequentemente representadas por números em uma escala, esses números possuem o significado de simplesmente indicar uma ordem. Assim, as variáveis são ordenadas em relação à quantia do atributo possuída. Podemos citar como exemplo uma escala Likert do tipo: quase nunca, raramente, às vezes, frequentemente e quase sempre. As escalas intervalares e escalas de razão (ambas métricas) fornecem um nível mais alto de precisão de medida, permitindo que quase todas as operações matemáticas sejam realizadas. As duas escalas têm unidades constantes de medida. Assim, diferenças entre quaisquer dois pontos em qualquer parte da escala são iguais. A única diferença entre as duas é que as escalas intervalares têm um ponto zero arbitrário, enquanto as escalas de razão têm um ponto zero absoluto. Um exemplo de escala intervalar são as escalas de temperatura. Um exemplo de escalas de razão são as balanças para medir pesos. As escalas de razão representam a maior precisão possível de medida, pois todas as operações matemáticas são possíveis com elas (HAIR *et al.*, 2005). Essa mesma nomenclatura é utilizada por Babbie (1999) para os níveis de mensuração das variáveis. No entanto, dependendo da fonte consultada, podem ocorrer pequenas variações quanto a esses nomes. A FIG. 2, além de sintetizar as informações sobre os níveis de mensuração, apresenta pequenas diferenças nos nomes:

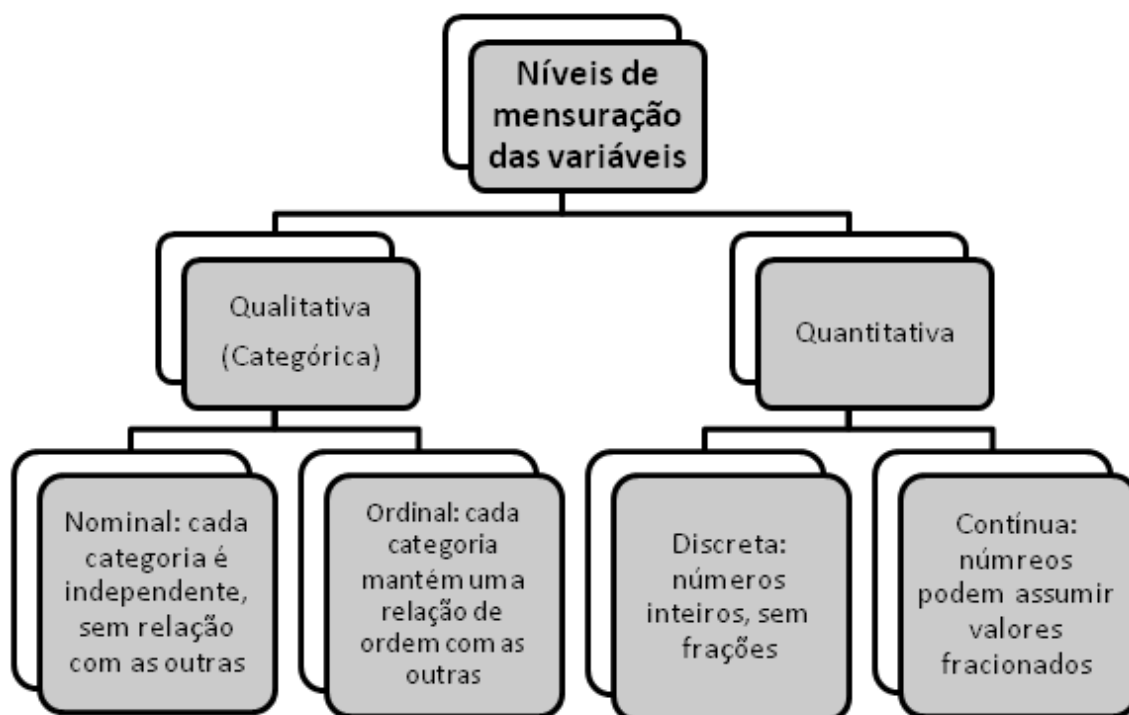


FIGURA 2 - Níveis de mensuração das variáveis

Como indicado na FIG. 2, as variáveis qualitativas são frequentemente denominadas de categóricas. As variáveis quantitativas também podem ser classificadas em discretas e contínuas. Além disso, outro termo é usado para a classificação de um tipo especial de variável categórica: variável dicotômica ou binária (*dummy variable*). Como o próprio nome indica, só existem duas categorias (exemplo: escola privada e pública).

O pesquisador precisa entender os diversos níveis de medida por duas razões. Primeiro, ao identificar a escala de medida de cada variável que está sendo utilizada, devemos estar atentos para que dados qualitativos não sejam erroneamente usados como quantitativos e vice-versa. Segundo, o nível de medida é crucial para determinar que técnicas multivariadas são as mais adequadas aos dados, considerando tanto variáveis independentes como dependentes (HAIR *et al.*, 2005). Como exemplo, podemos citar a regressão. Nessa técnica, o que determina o tipo de regressão a ser executada é a variável dependente: se a variável dependente for contínua é conduzida uma regressão linear, se for dicotômica uma regressão logística e assim por diante.

Outro exemplo pode ser dado com relação à nossa pesquisa. Muitos métodos estatísticos se baseiam na utilização da correlação entre os dados para estimar seus resultados (a regressão é um deles). A correlação mais conhecida é a de Pearson. Esse tipo de correlação requer que as variáveis tenham um nível de medida quantitativo. No

entanto, existem outros tipos de correlação: bisserial, policórica, polisserial e tetracórica. Fundamentalmente, essas correlações são consequência de combinações de tipos de variáveis diferentes (binárias, ordinais, quantitativas). Na análise das concepções de avaliação, utilizamos correlação policórica (quando as variáveis são medidas ordinais com três ou mais categorias) porque 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 (ou seja, uma variável ordinal). Na análise das definições de avaliação, utilizamos correlação tetracórica (quando as variáveis são medidas binárias) porque as respostas da parte complementar do questionário foram consideradas de forma dicotômica (HAIR *et al.*, 2005).

5.2.2 - Escalonamento Multidimensional

O Escalonamento Multidimensional (EM) (*Multidimensional Scaling - MDS*), ao examinar todos os pares de respostas, determina a proximidade relativa de um elemento com os outros e constrói um mapa perceptual (também denominado mapa espacial) das posições relativas para cada elemento. Elementos que estão próximos uns dos outros são claramente similares nas mentes dos respondentes. EM é um procedimento exploratório porque o pesquisador precisa interpretar o mapa perceptual e as dimensões subjacentes à posição relativa dos elementos. O número de dimensões subjacentes determina a forma geométrica do mapa perceptual. Uma dimensão cria uma linha, duas dimensões criam um plano e três dimensões criam um cubo. Embora seja possível ocorrer mais de quatro dimensões em uma solução de EM, o número de itens disponíveis para a classificação é uma restrição ao número de dimensões possíveis. Existe uma recomendação de pelo menos quatro objetos para cada dimensão; ou seja, se você quer duas dimensões você precisa de pelo menos oito objetos e inversamente se você tem somente 12 itens não deveria esperar mais do que três dimensões (STALANS, 1995).

Segundo Hair *et al.* (2005), o Escalonamento Multidimensional (também conhecido como mapeamento perceptual) se refere a várias técnicas que auxiliam o pesquisador a identificar dimensões chave inerentes a avaliações feitas por sujeitos quanto a objetos. As técnicas de EM podem inferir as dimensões subjacentes de diversos julgamentos de similaridades ou preferência fornecidos pelos respondentes com relação aos objetos. Dessa forma, o pesquisador pode determinar a imagem relativa percebida de um conjunto de objetos. O objetivo deve ser obter o melhor ajuste com o menor número possível de dimensões. A interpretação de soluções obtidas em mais de

três dimensões é muito difícil, sendo que geralmente não vale a pena mesmo que aconteça uma melhora no ajuste.

5.2.3 - Análise Fatorial

A Análise Fatorial (AF) é usada para investigar os padrões ou relações latentes para um número grande de variáveis e determinar se a informação pode ser resumida a um conjunto menor de fatores. O fator pode ser definido como uma combinação linear das variáveis originais. Os fatores representam as dimensões latentes (constructos) que resumem o conjunto original de variáveis, mantendo a representatividade das características das variáveis originais. Assim, os dois principais usos da análise fatorial são resumo e redução dos dados, que podem ser muito úteis à medida que o número de variáveis utilizadas em técnicas multivariadas aumenta. Além disso, a análise fatorial é diferente de técnicas de dependência como a regressão múltipla, nas quais uma variável é considerada como dependente e as outras independentes. A AF é uma técnica de interdependência, na qual todas as variáveis são consideradas ao mesmo tempo. Cada variável é prevista por todas as outras. Portanto, técnicas de dependência visam previsão e de interdependência visam identificação de estrutura (HAIR *et al.*, 2005).

As técnicas fatoriais podem atingir seus objetivos por uma perspectiva exploratória (Análise Fatorial Exploratória - AFE) ou por uma perspectiva confirmatória (Análise Fatorial Confirmatória - AFC). Bryant e Yarnold (2000) apontam claramente a diferença entre os dois métodos: na AFE, o pesquisador deixa os dados observados determinarem o modelo fatorial subjacente a posteriori (raciocínio indutivo para inferir um modelo a partir dos dados observados). Em contraste, na AFC o pesquisador deriva um modelo fatorial a priori (raciocínio dedutivo para fazer hipóteses de uma estrutura antecipadamente).

Uma técnica exploratória “deixa os dados falarem por si só”, não existe uma intervenção nossa como pré-determinar uma estrutura. Nós confiamos puramente na empiria dos dados e não estabelecemos restrições sobre a estimação ou número de componentes. No entanto, em outras situações o pesquisador tem ideias sobre a estrutura dos dados, baseado em algum referencial teórico e/ou pesquisas anteriores. Nesses casos, esperamos que a AF desempenhe um papel confirmatório. A Análise Fatorial Confirmatória é capaz de determinar como um modelo que está sendo testado se ajusta aos dados (KLEM, 2000). Dessa forma, é possível testar uma hipótese teórica por meio dos dados. Na AFC o pesquisador precisa especificar previamente toda a

estrutura a ser testada. Para isso, a pesquisa precisa ser dirigida pela teoria e hipóteses, onde a estatística é usada apenas como uma ferramenta.

Nessa pesquisa, utilizamos prioritariamente a Análise Fatorial Confirmatória. Ela é um procedimento melhor do que técnicas como a consistência interna (coeficiente alfa de Cronbach) 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). Dessa forma, esse procedimento é uma estimativa mais robusta para a validação de um questionário como o SCoA.

5.2.4 - Modelagem de Equações Estruturais

Técnicas estatísticas multivariadas como a regressão múltipla e a análise multivariada de variância compartilham uma mesma limitação: examinam apenas uma relação por vez. Mesmo técnicas que permitem múltiplas variáveis dependentes, ainda representam somente uma relação entre as variáveis independentes e dependentes. No entanto, frequentemente o pesquisador se depara com um conjunto de questões inter-relacionadas. Por isso, utilizamos a técnica de Modelagem de Equações Estruturais (MEE) (*Structural Equation Modeling - SEM*), uma extensão de várias técnicas multivariadas. A MEE examina uma série de relações de dependência simultaneamente, sendo muito útil quando uma variável dependente se torna independente em relações de dependência subsequentes. Esse conjunto de relações é a base da MEE. Portanto, a Modelagem de Equações Estruturais pode ser definida como uma técnica multivariada que combina aspectos da regressão múltipla (relações de dependência) e análise fatorial (representando fatores) para estimar diversas relações de dependência inter-relacionadas simultaneamente (HAIR *et al.*, 2005).

A análise fatorial lida com conceitos e não permite a especificação de relações causais. A técnica conhecida como Análise de Caminhos (*Path Analysis*) permite a especificação de relações causais (inclusive de várias relações de dependência simultaneamente), mas trabalha somente com variáveis observáveis. Assim, a solução para essa questão é a técnica de Modelagem de Equações Estruturais. De uma maneira geral, a MEE também pode ser vista como uma combinação de Análise de Caminhos (utilização de regressão) e Análise Fatorial. Na Modelagem de Equações Estruturais, a preocupação é com a ordem preditiva dos fatores, que são variáveis latentes que se relacionam com variáveis medidas. A MEE junta a sofisticação de medida dos fatores com modelos que especificam relações causais (KLEM, 2000).

Finalmente, destacamos a diferença entre modelos de medida e modelos estruturais. Modelos de medida explicam a relação entre estruturas dentro de um fenômeno. A estrutura das concepções de avaliação, medida nessa pesquisa por meio de análise fatorial, é um exemplo de modelo de medida. Já os modelos estruturais predizem relações entre diferentes estruturas (HOYLE, 1995). Como exemplo, podemos citar um dos objetivos específicos do presente trabalho: analisar como as concepções de avaliação predizem as definições de avaliação dos alunos (relações entre concepções e definições de avaliação).

5.2.5 - Qualidade de ajuste: índices de ajuste do modelo

Os métodos estatísticos nos oferecem uma série de indicadores (índices de ajuste) para saber se nossos resultados são satisfatórios ou não. Primeiramente, quanto ao Escalonamento Multidimensional, são considerados indicadores de um bom ajuste quando o valor do *Kruskal's strêss* é $<.15$ e quando o coeficiente de determinação (R^2), a proporção da variância explicada pelas dimensões, é $>.90$ (STALANS, 1995). Já Hair *et al.* (2005) argumentam que valores do $R^2 >.60$ são considerados aceitáveis. Na presente pesquisa, utilizamos o critério proposto por Stalans.

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 *gamma hat*, 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 $CFI \geq .90$, $TLI \geq .90$, $\gamma \hat{} \geq .90$, $RMSEA \leq .08$ 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; HAU; WEN, 2004; MARSH; HAU; GRAYSON, 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 de especificação do modelo muito pequenos se o tamanho da amostra é

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; RENSVOOLD, 2002). Já os índices RMSEA, SRMR (*badness-of-fit índices*) e *gamma hat* (*goodness-of-fit indice*) 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.

5.3 - Análise de conteúdo

Utilizamos a análise de conteúdo para investigar as questões abertas que integram a parte complementar do SCoA (ANEXO A). A análise de conteúdo está ligada a uma concepção crítica e dinâmica de linguagem, tendo como ponto de partida uma mensagem verbal ou escrita que está vinculada a determinadas condições contextuais de produção. Seu campo de aplicação é grande, pois considera que qualquer comunicação pode ser interpretada por suas técnicas. Assim, a mensagem é tratada com a intenção de inferir conhecimentos sobre o emissor e seu meio (BARDIN, 1977).

Gall, Gall e Borg (2007) recomendam os seguintes passos na análise de conteúdo: 1) especificar as questões de pesquisa, hipóteses ou objetivos; 2) selecionar uma amostra de documentos para analisar; 3) desenvolver um procedimento de categorização-codificação; 4) codificar os dados; 5) interpretar os resultados.

Dessa forma, utilizamos a análise de conteúdo para interpretar as respostas dos alunos nas questões abertas (análise de texto). Esse foi o método qualitativo adotado no presente trabalho (a análise de conteúdo também pode ser usada dentro de uma perspectiva mais quantitativa).

5.4 - 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, Comercio Exterior, Engenharia Civil, Farmácia, Física, Gestão Comercial, Gestão de Recursos Humanos, Medicina Veterinária, Nutrição, Pedagogia, Processos Gerenciais, Psicologia, Serviço Social, Terapia Ocupacional.

Apesar da amostragem dessa pesquisa ser não-probabilística, tentamos garantir a maior heterogeneidade possível investigando cursos de todas as áreas (Humanas, Exatas e Biológicas). A diferença do número de alunos das duas IES se deveu à dificuldade de acesso na Universidade Federal, onde a autorização para a realização da pesquisa foi mais problemática.

Quanto ao número de homens e mulheres da amostra, a TAB. 1 mostra dados do Censo da Educação Superior de 2008 sobre as matrículas em cursos de graduação de acordo com o sexo.

TABELA 1
Matrículas em Cursos de Graduação Presenciais: sexo dos matriculados em 2008

	Universidade Federal			Centro Universitário privado		
	<i>Total</i>	<i>Masculino</i>	<i>Feminino</i>	<i>Total</i>	<i>Masculino</i>	<i>Feminino</i>
Brasil	600.772	297.143	303.629	384.896	171.525	213.371
Sudeste	166.087	81.558	84.529	220.565	98.982	121.583
Minas Gerais	76.809	38.853	37.956	32.888	14.398	18.490

Fonte: MEC/INEP/DEED (Censo da Educação Superior de 2008)

Como indicado na TAB. 1, o número de homens e mulheres matriculados em Universidades Federais é próximo no Brasil, no Sudeste e especialmente em Minas Gerais. Já nos Centros Universitários, existe uma diferença considerável no número de matrículas quando tomamos como referência o Brasil e a região sudeste. Os valores se tornam mais ou menos próximos quando consideramos o estado de Minas Gerais.

Considerando que nossa amostra é composta em sua maioria por alunos do Centro Universitário privado, ela reproduziu parcialmente os dados do censo: acompanhou uma tendência apresentada no Brasil e Sudeste, mas não acompanhou os números de Minas Gerais. Independente disso, pelo fato da nossa amostra não ser aleatória, não podemos generalizar os resultados da pesquisa.

5.5 - Etapas da pesquisa

Essa pesquisa foi composta por vários estudos, que se caracterizaram pela utilização de técnicas múltiplas. A FIG. 3 apresenta os estudos que foram desenvolvidos:

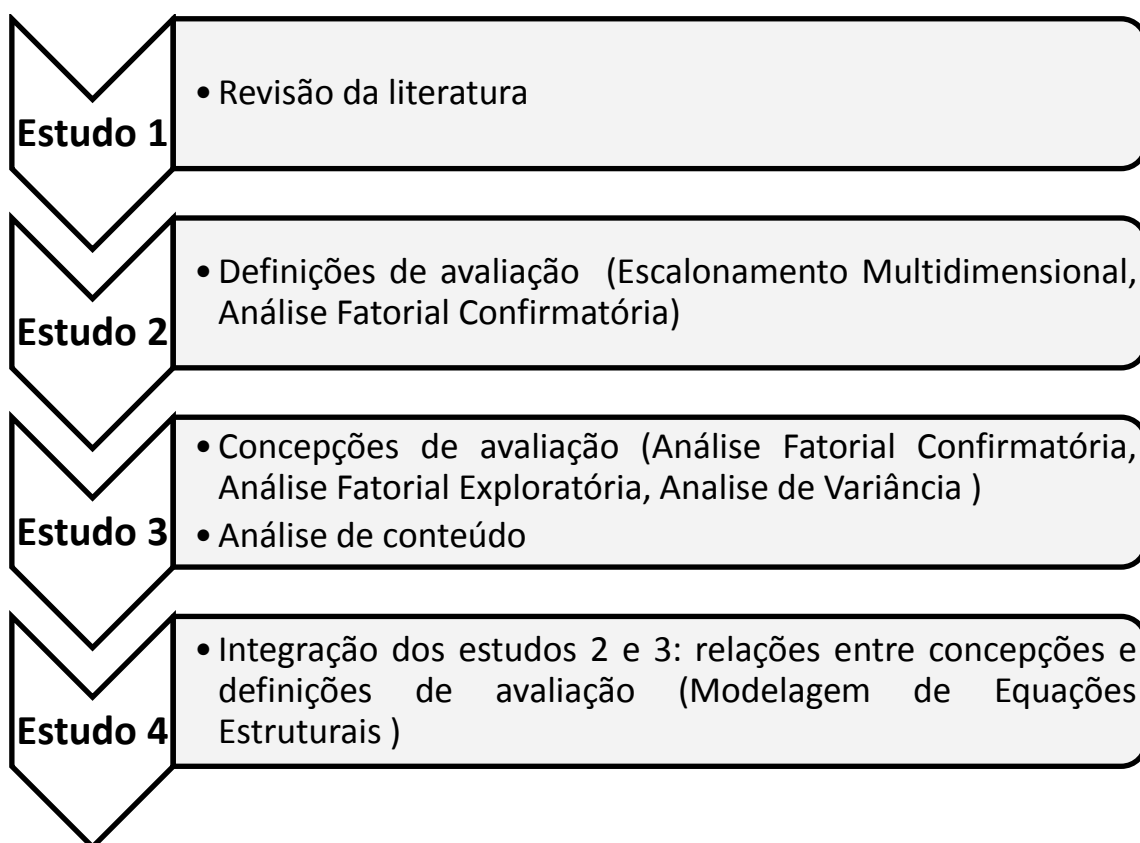


FIGURA 3 - Esquema dos estudos desenvolvidos na pesquisa

Como indicado na FIG. 3, foram cumpridas diversas etapas na pesquisa. No estudo 1, realizamos uma revisão da literatura sobre: avaliação no ensino superior brasileiro (reformas na área da avaliação, ensino público e privado, legislação educacional, SINAES); teoria sobre concepções de avaliação e concepções de avaliação dos alunos universitários brasileiros; teoria ligada ao SCoA (aprendizagem auto-regulada) e diferenças culturais/pesquisas desenvolvidas na Nova Zelândia. Todas essas temáticas geraram as hipóteses do presente trabalho. No estudo 2, investigamos as definições de avaliação dos alunos usando as técnicas estatísticas denominadas Escalonamento Multidimensional e Análise Fatorial Confirmatória. No estudo 3, analisamos as concepções de avaliação dos alunos por meio de Análise Fatorial

Confirmatória e Exploratória, Análise de Variância (ANOVA). Esse último método foi utilizado com o objetivo de comparar as concepções de avaliação dos alunos pertencentes a uma IES pública e a uma IES privada. Ainda no estudo 3, utilizamos um método qualitativo: análise de conteúdo para investigar as questões abertas do SCoA. Os estudos 2 e 3 também incluíram a adaptação e validação do SCoA. Já o estudo 4 visou uma integração dos estudos anteriores por meio do estabelecimento de relações entre concepções e definições de avaliação dos estudantes. Nesse estudo, como tínhamos o objetivo de responder uma questão causal, foi empregada a técnica de Modelagem de Equações Estruturais.

Passamos a descrever de forma mais detalhada alguns procedimentos metodológicos que utilizamos.

5.5.1 - Procedimentos teóricos

Usamos nessa pesquisa o modelo de validação de instrumentos proposto por Pasquali (1999), que foi discutido na seção 3.4.

Inicialmente, fizemos uma tradução da versão em inglês do SCoA para o português (ANEXO B). Participaram dessa tradução o pesquisador, o orientador da pesquisa e um professor de inglês. O grupo trabalhou em conjunto na tradução. Depois, realizamos o procedimento de tradução inversa (*back translation*) (ANEXO B). Foi feita uma tradução do português para o inglês por três pesquisadores que não se envolveram na tradução original. Eles realizaram esse trabalho de maneira independente e não tiveram contato uns com os outros. A partir da análise desse conjunto de traduções, foi possível obter um consenso sobre a tradução do questionário. Participaram dessa etapa final da tradução o pesquisador e o orientador da pesquisa. O ANEXO B mostra todos os passos envolvidos no processo de tradução.

Posteriormente, realizamos uma análise teórica dos itens. Doze alunos participaram de entrevistas semi-estruturadas¹⁸ (ANEXO C) em grupos: cinco alunos da Universidade Federal e sete alunos do Centro Universitário privado. Na entrevista, pedimos aos integrantes do grupo que respondessem ao SCoA individualmente e em silêncio. Eles foram instruídos a marcar a lápis qualquer dúvida que tivessem sobre o questionário durante a leitura. Depois, todas as dúvidas dos estudantes foram discutidas em grupo. Utilizando essas entrevistas, pequenas alterações foram realizadas no SCoA

¹⁸ Na entrevista semi-estruturada, o entrevistador possui um roteiro de perguntas que serve para guiar a entrevista. Entretanto, o entrevistador também tem liberdade para fazer outros questionamentos que julgar necessário.

(ver capítulo 6). Assim, no final dessa fase de validação, algumas modificações foram incluídas no questionário.

5.5.2 - Procedimentos empíricos (experimentais)

Após a etapa de validação teórica, foi realizado um estudo piloto em uma sala de aula da Universidade Federal com 37 alunos e em uma sala do Centro Universitário privado com 33 alunos. Utilizamos como instrução oral a primeira página que elaboramos para o SCoA, que era lida pelo aplicador. Optamos por criar essas instruções para melhorar tanto a aplicação do SCoA quanto as informações coletadas por esse instrumento. Já apontamos a importância desse procedimento ao discutir a aplicação em grupo auto-administrada na seção 5.1 (maior estruturação da aplicação em grupos, criação de um ambiente similar). Após o estudo 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. Esse foi um passo importante na preparação do questionário.

Depois, foi feito um contato com diversos setores das duas IES solicitando a aprovação da pesquisa (ANEXO D). O SCoA foi aplicado em uma amostra de 756 alunos. A aplicação foi realizada de forma coletiva na sala de aula dos estudantes. O professor cedia um tempo da sua aula (geralmente, de vinte a trinta minutos). Alguns professores permaneciam na sala durante a aplicação, outros não. Isso não influenciou os resultados da pesquisa, pois a presença do professor não era necessária durante a aplicação do questionário. O SCoA foi aplicado pessoalmente pelo pesquisador e por uma colaboradora da pesquisa (na época, uma aluna de Mestrado em Educação). Essa colaboradora recebeu um treinamento dado pelo pesquisador para garantir que as condições de aplicação fossem as mesmas. Durante a própria aplicação, obtivemos a autorização dos alunos para participar da pesquisa por meio da assinatura do termo de consentimento livre e esclarecido¹⁹ (ANEXO E). Assim, os estudantes que aceitavam participar entregavam o questionário juntamente com o termo de consentimento assinado. Quanto aos procedimentos analíticos (estatísticos) utilizados, eles já foram abordados anteriormente nesse capítulo.

¹⁹ O presente trabalho foi submetido ao Comitê de Ética em Pesquisa (COEP) da Universidade Federal de Minas Gerais (UFMG) e seguiu todas as normas e diretrizes contidas na Resolução 196/96, que regulamenta as pesquisas envolvendo seres humanos.

CAPÍTULO 6 - RESULTADOS E DISCUSSÃO

6.1 - Adaptação e validação do SCoA

Como consequência dos procedimentos teóricos, empíricos (experimentais) e analíticos (estatísticos) que utilizamos, algumas modificações foram realizadas no SCoA. São elas:

Instruções

Como já mencionamos anteriormente, elaboramos uma primeira página para o SCoA com instruções que eram lidas no momento da aplicação (ANEXO A). A versão original em inglês não possui uma primeira página com instruções. Incluímos nessa parte aspectos importantes como:

- Dados dos participantes.
- Delimitação do nível de ensino investigado: “Esse questionário contém afirmações sobre experiências relativas à avaliação. Para respondê-lo, leve em consideração as suas experiências de avaliação no Ensino Superior de um modo geral”.
- Delimitação do *setting* de aplicação, diferenciando de outras experiências escolares: “Não existem respostas “certas” ou “erradas”. Isso não é uma prova nem um teste. Responda calmamente sem se preocupar com o tempo”.
- Forma de resposta do questionário: “O questionário deve ser respondido individualmente e em silêncio”.
- Prevenção de problemas na aplicação, solicitando a conferência de algumas questões: “Esse questionário tem 4 folhas. Confira se o seu tem 4 folhas e responda a todas”.

Itens sobre as concepções de avaliação

Apenas uma modificação foi realizada nessa parte. O item 33 da versão original do SCoA (*Assessment tells my parents how much I've learnt* - em português: A avaliação diz para os meus pais o quanto eu aprendi) foi excluído da versão brasileira. Essa exclusão aconteceu por causa das entrevistas conduzidas com os alunos na parte de procedimentos teóricos da pesquisa. Os estudantes relataram que esse item não fazia sentido para eles no nível de ensino superior, sendo que isso se aplicaria quando eram mais novos.

Itens sobre as definições de avaliação

Cinco itens da “Parte complementar do questionário” foram modificados para que as atividades se adaptassem melhor ao contexto brasileiro (ANEXO B). A seguir apresentamos todas as modificações realizadas (nossos comentários estão entre parênteses).

Os itens da versão original em inglês excluídos foram:

- *An examination that takes one to three hours* - Uma prova que leva de uma a três horas.
- *The teacher grades me on a written test that he or she made up* - O professor me dá uma nota em um teste escrito que ele fez. (Esses dois primeiros itens sobre prova foram substituídos por itens mais específicos: prova com consulta, prova em dupla, prova com questões de múltipla escolha, prova com questões abertas)
- *The teacher grades me on a written test that was written by someone other than the teacher* - O professor me dá uma nota em um teste escrito que foi feito por alguém diferente do professor (No ensino superior brasileiro os próprios professores elaboram suas atividades avaliativas)
- *The teacher scores a portfolio of work I have done over the course of a term or school year* - O professor atribui uma nota a um portfólio de trabalhos que eu fiz durante um semestre ou ano escolar. (Considerando nossas leituras sobre portfólio, essa ainda parece ser uma prática avaliativa incipiente no ensino superior brasileiro. Além disso, nas entrevistas realizadas com os estudantes na parte de procedimentos teóricos nenhum deles sabia o que era portfólio).
- *The teacher uses a checklist to judge my in-class performance* - O professor usa um *checklist* para julgar o meu desempenho dentro da sala de aula. (Essa prática avaliativa não parece ter um paralelo no ensino superior brasileiro. Poderíamos pensar em algo similar no ensino básico).

Os itens novos da versão em português foram:

- O professor atribui uma nota ou avalia um trabalho em grupo
- O professor aplica prova com consulta
- O professor aplica prova em dupla
- O professor aplica prova com questões de múltipla escolha
- O professor aplica prova com questões abertas

Em resumo, a maioria desses itens detalhava algumas situações de prova. Destacamos que alguns desses tipos de prova parecem ser bem específicos da realidade brasileira como, por exemplo, a prova em dupla. Também incluímos uma prática avaliativa muito comum no ensino superior brasileiro: o trabalho em grupo.

6.2 - Definições de avaliação

Relatamos a seguir os resultados do SCoA que indicam como os estudantes definem a avaliação com referência à uma série de práticas avaliativas.

6.2.1 - Escalonamento Multidimensional

Como parte integrante do SCoA, os alunos tinham que responder a pergunta: “Quando você pensa na palavra AVALIAÇÃO, quais tipos de atividades avaliativas vêm a sua mente?” (ANEXO A). Os estudantes podiam escolher até 12 tipos de atividades avaliativas diferentes. A lista de atividades foi derivada da lista utilizada por Brown *et al.* (2009). Entretanto, cinco itens foram modificados na versão brasileira (eles estão marcados com um asterisco na TAB. 2). A instrução dessa parte do instrumento era “marque um X nas opções que quiser. Você pode marcar quantas considerar necessário.” As respostas foram consideradas de forma dicotômica (0= não selecionada; 1= selecionada). Um procedimento exploratório de Escalonamento Multidimensional (EM) foi usado para determinar as dimensões subjacentes para as respostas dos alunos aos 12 tipos de atividades avaliativas (FIG. 4). O *software* SPSS versão 17.0 foi utilizado para realizar o EM.

Um Escalonamento Multidimensional para duas, três e quatro soluções dimensionais foi computada usando distância euclidiana nas 12 medidas binárias. As soluções de três e quatro dimensões foram rejeitadas por causa do ajuste pobre dos itens às dimensões. A solução de duas dimensões teve um valor de Kruskal strêss de .08 e um R^2 de .97. A dimensão 1 (eixo horizontal) foi interpretada como o grau de formalidade das atividades avaliativas, enquanto a dimensão 2 (eixo vertical) foi entendida como o lócus de controle das várias atividades avaliativas (ou seja, o aluno ou o professor). Essa análise de duas dimensões criou um mapa perceptual 2 x 2: informal/controle do aluno, informal/controle do professor, formal/controle do aluno, formal/controle do professor. Assim, itens que caem completamente dentro de cada quadrante devem se enquadrar à interação das duas dimensões; enquanto itens que caem em cima de um eixo devem estar na metade do caminho entre os dois fins de cada dimensão. Destacamos que cinco elementos (A, C, E, I, J) tiveram valores da dimensão 2 próximos a zero (+/- .10),

enquanto nenhum dos valores da dimensão 1 foram tão próximos de zero. Tais resultados sugerem que os participantes tiveram pouca dificuldade para distinguir a dimensão de formalidade e mais dificuldade para discernir o lócus de controle. Realizamos agora uma análise mais detalhada de cada um dos quadrantes encontrados no EM.

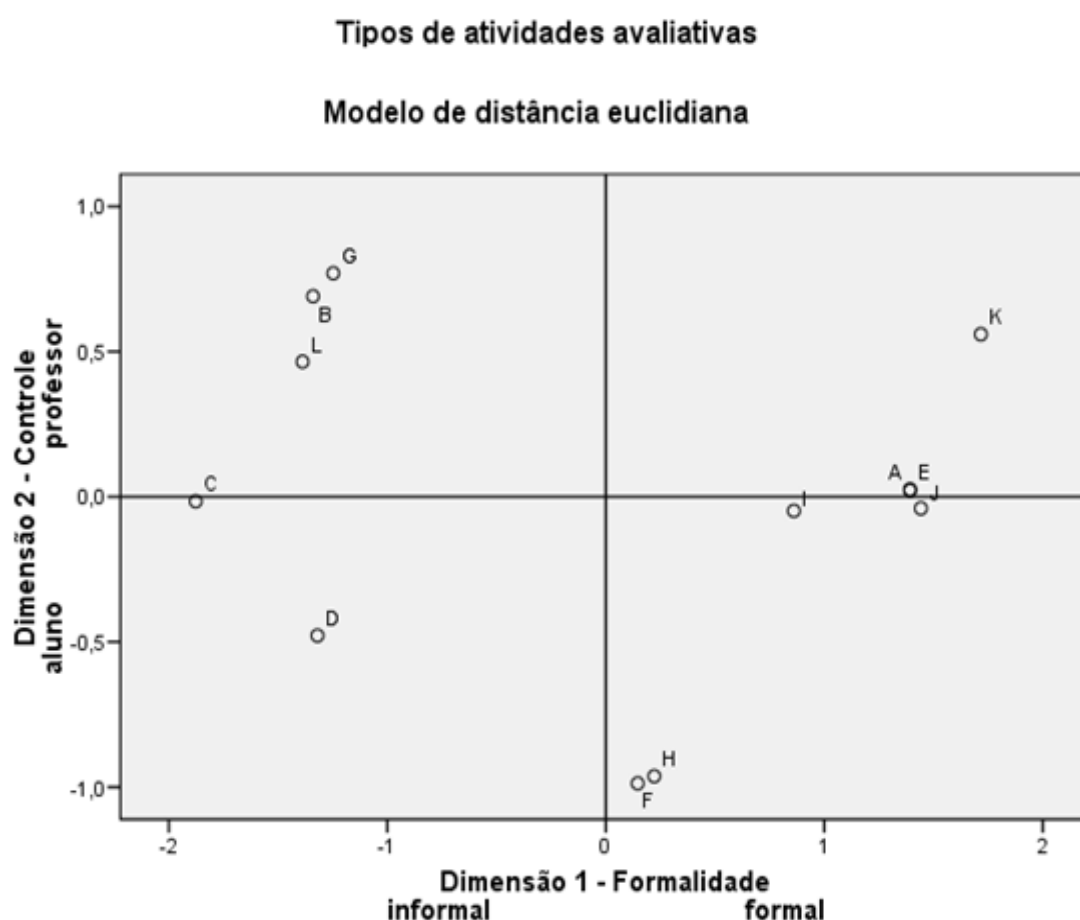


FIGURA 4 - Mapa perceptual: tipos de atividades avaliativas

Nota: A letra E está oculta pela letra A porque os seus valores são muito similares.

Formal/controle do professor

Os itens A, E, I, J, K caíram nesse grupo. Esses itens foram os mais selecionados pelos alunos, com o item k (o professor aplica prova com questões abertas) sendo selecionado por quase todos os participantes (87%). As outras atividades formais controladas pelo professor foram selecionadas por dois terços a três quartos dos estudantes. Portanto, as respostas dos alunos mostraram uma ênfase em métodos de avaliação convencionais que utilizam métodos formais controlados pelo professor.

TABELA 2
Tipos de atividades avaliativas e os quadrantes do Escalonamento Multidimensional

Tipos de atividades avaliativas	Frequência	Porcentagem	Dimensão 1	Dimensão 2
<i>Formal/controle do professor</i>				
*K- O professor aplica prova com questões abertas	610	87	1.72	0.56
*A- O professor atribui uma nota ou avalia um trabalho em grupo	542	77	1.39	0.02
E- O professor atribui uma nota ou avalia um trabalho escrito que eu entrego	533	76	1.39	0.03
*J- O professor aplica prova com questões de múltipla escolha	525	75	1.44	-0.04
I- O professor me dá uma nota em um trabalho feito na sala de aula	448	64	0.86	-0.05
<i>Média</i>	531.6	75.8	1.36	0.10
<i>Informal/controle do professor</i>				
G- O professor me observa na sala de aula e julga o meu aprendizado	260	37	-1.25	0.77
B- Eu atribuo uma nota ou avalio meu próprio desempenho (auto-avaliação)	227	32	-1.34	0.69
L- O professor atribui uma nota ao meu desempenho depois de se reunir ou discutir comigo sobre o meu trabalho	210	30	-1.39	0.47
C- Meus colegas atribuem uma nota ou avaliam meu desempenho	92	13	-1.88	-0.02
<i>Média</i>	197.25	28	-1.47	0.48
<i>Formal/controle do aluno</i>				
*H- O professor aplica prova em dupla	321	46	0.22	-0.96
*F- O professor aplica prova com consulta	318	45	0.15	-0.99
<i>Média</i>	319.5	45.5	0.19	-0.98
<i>Informal/controle do aluno</i>				
D- O professor me faz perguntas em voz alta na sala de aula	187	27	-1.32	-0.48
<i>Média</i>	187	27	-1.32	-0.48

Nota: os itens com * são diferentes da versão original do SCoA. Eles foram criados para a versão brasileira do instrumento.

Informal/controle do professor

Esse quadrante descreve práticas avaliativas informais que foram controladas pelo professor. Itens G (observação do professor), B (auto-avaliação) e L (professor atribui nota depois de se reunir) caíram nesse grupo e foram selecionados por aproximadamente um terço dos alunos (30 a 37%). A primeira vista, o item B deveria

ser uma atividade controlada pelo estudante ao invés de ser uma atividade controlada pelo professor. Existem algumas explicações possíveis para a proximidade da autoavaliação com os itens L e G. Em primeiro lugar, pode ser simplesmente que os alunos se avaliam de acordo com um direcionamento do professor (ex: um relatório escrito ou uma discussão oral), significando que o professor está no controle do processo. Em uma segunda possível explicação, parece que poucos professores do ensino superior atribuem pontos (ou até mesmo usam) para a autoavaliação dos alunos, sugerindo que a avaliação real ainda é controlada pelo professor. Em terceiro lugar, se a autoavaliação visa se aproximar à avaliação do professor (FALCHIKOV; BOUD, 1989), então não deveria ser uma surpresa considerar que mesmo a autoavaliação ainda é controlada pelo professor.

O lugar do item C (avaliação dos colegas) foi problemático porque o valor na dimensão 2 foi apenas um pouco menos do que zero. É possível considerar que a avaliação dos colegas é controlada pelos alunos. Entretanto, se a avaliação dos colegas no ensino superior funciona como a autoavaliação (item B), porque é feita de acordo com o direcionamento do professor ou tenta ser aproximar aos critérios do professor (FALCHIKOV; GOLDFINCH, 2000), então o item C parece um forte candidato para a inclusão nesse quadrante. A correlação do item C com os outros três membros desse quadrante foi estatisticamente significativa com uma média de $r=.22$ ($DP=.04$); enquanto a correlação com o único item de “informal/controle do aluno” foi somente $r=.14$. Assim, concluímos que esse item é uma atividade informal controlada pelo professor para a maioria dos estudantes.

Formal/controle do aluno

Esse quadrante é formado por dois itens que não tem uma relação prévia nas pesquisas neozelandesas. O item H (prova em dupla) e o item F (prova com consulta) foram selecionados por quase metade dos estudantes e são claramente avaliações formais nas quais é permitido um controle considerável aos alunos sobre o que acontece. Os estudantes podem trabalhar juntos e consultar uns aos outros mesmo sob condições de exame.

Informal/controle do aluno

Somente um item (D - professor faz perguntas em voz alta na sala) ficou nesse quadrante e foi selecionado por um quarto dos participantes. À primeira vista, isso parece ser uma localização inesperada para uma avaliação que é claramente iniciada

pelo professor. Entretanto, essa localização se torna mais lógica quando consideramos o fato de que os alunos determinam como e em que extensão eles participam nessa atividade.

Esperávamos a partir do trabalho de Brown *et al.* (2009) realizado com estudantes neozelandeses do ensino médio encontrar dois grupos de práticas avaliativas (hipótese 2). Tal hipótese encontrou apoio na primeira dimensão de formalidade. Encontramos uma distinção clara entre práticas avaliativas formais e interativo-informais. Entretanto, uma dimensão adicional (lôcus de controle) foi necessária para compreender totalmente as definições de avaliação dos alunos do ensino superior (o controle dos estudantes sobre o processo avaliativo muda a forma como a avaliação é experienciada e compreendida). Assim, o presente estudo indicou que duas dimensões (formalidade e lôcus de controle) e a sua interação poderiam explicar como os estudantes definem a avaliação. Nossa pesquisa resultou em quatro grupos de atividades avaliativas: informal/controle do aluno, informal/controle do professor, formal/controle do aluno, formal/controle do professor. Ao comparar esses resultados com as pesquisas neozelandesas, podemos afirmar que eles foram um refinamento na compreensão das definições de avaliação dos estudantes. Consideramos que existem duas principais explicações para esse resultado (além das próprias diferenças culturais entre os dois países). Primeiro, cinco dos 12 itens do SCoA foram modificados para uma melhor adaptação à realidade brasileira. Segundo, foram pesquisados níveis de ensino diferentes (médio e superior).

A revisão da literatura brasileira que realizamos e as pesquisas de Brown *et al.*, (2009) e Brown, Irving, e Peterson, (2008) nos levaram a esperar que a avaliação fosse definida principalmente como métodos formais/convencionais do tipo exames e testes (hipótese 3). Essa hipótese se confirmou, pois as respostas dos alunos mostraram uma ênfase em métodos de avaliação convencionais. Além disso, nove dos 12 tipos de atividades avaliativas caíram em quadrantes controlados pelo professor. Até mesmo práticas supostamente controladas pelos alunos (autoavaliação, avaliação dos colegas) foram percebidas como práticas controladas pelo professor.

Essa dominância de avaliação controlada pelo professor pode refletir a necessidade ou desejo dos alunos de um forte direcionamento por parte dos professores (PETERSON; IRVING, 2008). Isso também poderia refletir tentativas dos estudantes de replicar critérios, julgamentos dos professores (FALCHIKOV; BOUD, 1989; FALCHIKOV; GOLDFINCH, 2000) ou simplesmente os alunos percebem as práticas

interativo-informais como irrelevantes, considerando que os professores não pontuam com frequência tais práticas.

6.2.2 - Análise Fatorial Confirmatória

Ainda com relação às definições de avaliação dos alunos (dimensões subjacentes nas respostas da parte complementar do SCoA - 12 tipos de atividades avaliativas), analisamos os índices de ajuste de três modelos: o modelo de quatro fatores encontrado no procedimento exploratório do Escalonamento Multidimensional (Kruskal stress= .08; $R^2 = .97$) - informal/controle do aluno, informal/controle do professor, formal/controle do aluno, formal/controle do professor; um modelo alternativo de três fatores - informal/controle do professor, formal/controle do aluno, formal/controle do professor; o modelo de dois fatores encontrado por Brown *et al.* (2009) - formal e informal.

O *software* MPLUS versão 3.11 foi utilizado para realizar a Análise Fatorial Confirmatória (AFC) do SCoA. 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 usando um estimador robusto do MPLUS para a análise de indicadores categóricos denominado *Mean and variance-adjusted weighted least squares estimator*²⁰ (WLSMV). Nesse ponto, fazemos uma crítica ao trabalho do autor original do SCoA. Brown utilizou em suas análises, até o presente momento, o programa AMOS (pacote complementar do SPSS) que não possui estimadores robustos para variáveis categóricas e usa correlação de Pearson para as análises. Assim, o MPLUS apresenta uma série de vantagens sobre outros programas e é considerado um dos melhores *softwares* atualmente, principalmente para variáveis latentes.

A fim de avaliar o ajuste do modelo, a estatística qui-quadrado, o índice de ajuste comparativo (CFI), o índice de Tucker-Lewis (TLI), o índice *gamma hat*, a raiz do erro quadrático médio de aproximação (RMSEA) e a raiz do resíduo quadrático médio padronizado (SRMR) foram avaliados (FAN; SIVO, 2007).

A solução com quatro fatores encontrada no EM foi inadmissível na AFC. A TAB. 3 indica os valores encontrados para as respostas ao SCoA brasileiro.

²⁰ Nesse trabalho, foi utilizado para a tradução dos termos estatísticos um glossário elaborado por Ravagnani e Catelan (2002) e livros de estatística em português. Não foram encontradas, entretanto, as traduções para todas as expressões. Por isso, alguns termos são apresentados em inglês.

TABELA 3
 Comparação de índices de ajuste dos modelos da AFC: Definições de avaliação

Modelos	χ^2	gl	CFI	TLI	Gamma hat	RMSEA	SRMR
2 fatores	205.489	36	.91	.91	.96	.079	.105
3 fatores	176.762	34	.92	.92	.97	.075	.100

Nota: Todos os valores χ^2 foram estatisticamente significantes ($p < .0000$).

O valor do qui-quadrado encontrado por meio do estimador robusto *Mean and variance-adjusted weighted least squares estimator* (WLSMV) não pode ser usado para testes de diferença de qui-quadrado da maneira convencional. Nesse caso, o manual do programa MPLUS recomenda rodar um teste de diferença de qui-quadrado ajustado (MUTHÉN; MUTHÉN, 2006). O resultado do teste foi ($\chi^2 = 27.184$, $p < .0000$), indicando que o modelo com três fatores melhora o ajuste do modelo de maneira estatisticamente significativa quando comparado ao modelo de dois fatores. Além disso, acreditamos que o modelo de três fatores é mais consistente do ponto de vista teórico. Ele permite captar melhor a complexidade de algumas atividades avaliativas do ensino superior brasileiro (ex: o professor aplica prova em dupla).

A estatística qui-quadrado do modelo de três fatores foi significativa. Entretanto, como já discutimos, a estatística qui-quadrado é sensível a erros de especificação do modelo muito pequenos se o tamanho da amostra é 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 muito bem à estrutura fatorial de três fatores.

Aqui, uma questão precisa ser explicitada: encontrar um bom ajuste do modelo como esse, não significa provar empiricamente que esse modelo está “correto”. Ou seja, existe a possibilidade de ajustar modelos diferentes com índices de valores muito próximos ou iguais. Essa ponderação sobre o ajuste estatístico de um modelo vale para todos os resultados do presente trabalho. Tal ponderação evidencia a importância de escolher variáveis e modelos por meio de teoria e pesquisas anteriores.

As cargas fatoriais padronizadas de todos os itens relacionados às definições de avaliação do SCoA brasileiro são apresentadas na TAB. 4.

TABELA 4
Cargas fatoriais padronizadas dos itens do SCoA Brasileiro: Definições de avaliação

<i>Informal controle</i>		<i>Formal controle</i>		<i>Formal controle</i>	
<i>professor</i>		<i>aluno</i>		<i>professor</i>	
Item	Carga	Item	Carga	Item	Carga
B	.59	F	.84	A	.70
C	.68	H	.98	E	.76
D	.47			I	.74
G	.60			J	.70
L	.66			K	.60

Como indicado na TAB. 4, a maioria dos itens teve cargas fatoriais altas (especialmente os itens do fator Formal/controle do aluno). Esses resultados evidenciam também um ponto passível de crítica que merece ser pensado em pesquisas futuras. Após a AFC, o fator “Formal/controle do aluno” ficou somente com dois itens. De um ponto de vista técnico, não é muito adequado ter um fator só com dois itens. Por isso, recomendamos que novos itens sejam incluídos nesse fator em futuras análises.

Quanto à consistência interna, os fatores apresentaram valores variando entre moderado e bom (Informal/controle do professor, Cronbach $\alpha = .55$; Formal/controle do aluno, Cronbach $\alpha = .76$; Formal/controle do professor, Cronbach $\alpha = .65$). No entanto, como já destacamos anteriormente, a AFC proporciona um indicador mais robusto da qualidade do instrumento porque leva em consideração ao mesmo tempo correlações, covariâncias e resíduos.

As correlações entre os fatores foram: “informal/controle professor” e “formal/controle professor”, $r = .23$; “informal/controle professor” e “formal/controle aluno”, $r = .14$; “formal/controle aluno” e “formal/controle professor”, $r = .81$. Os valores baixos das correlações entre os fatores “formal” e “informal” indicam que essas duas dimensões são claramente distintas. Também merece destaque o fato das duas correlações serem positivas. Já a correlação entre os fatores “formal/controle aluno” e “formal/controle professor” foi alta. Esse valor poderia ser esperado porque os dois fatores compartilham o aspecto da formalidade da avaliação.

6.3 - Concepções de avaliação

Relatamos a seguir os resultados do SCoA sobre as concepções de avaliação.

6.3.1 - Consistência interna e evidências de validade

A literatura sobre testes educacionais e psicológicos recomenda o cálculo da consistência interna como um dos métodos de validação de instrumentos (FRASER, 1998). O coeficiente de consistência interna mede a similaridade entre os itens do teste levando em consideração a correlação deles. A TAB. 5 compara a consistência interna do SCoA no Brasil com os resultados de Brown *et al.* (2009) na Nova Zelândia.

TABELA 5
Comparação da consistência interna do SCoA entre amostras do Brasil e Nova Zelândia

Fator	Coeficiente Alfa de Cronbach	
	<i>Brasil</i>	<i>Nova Zelândia</i>
Melhora (aluno)	.83	.88
Melhora (professor)	.83	.80
Emoção (aluno)	.67	.75
Emoção (sala)	.86	.87
Irrelevante (ruim)	.52	.77
Irrelevante (ignorar)	.59	.67
Responsabilização (aluno)	.72	.71
Responsabilização (escola)	.76	.63

Nota: tamanho das amostras: Brasil (756), Nova Zelândia (705).

Como apontado na TAB. 5, o valor do Coeficiente Alfa de Cronbach para cada fator variou entre .52 e .86 no Brasil e entre .63 e .88 na Nova Zelândia. De acordo com Hair *et al.* (2005), os valores de Alfa de Cronbach de .60 a .70 são considerados o limite inferior da aceitabilidade. Na presente pesquisa, utilizamos o critério de corte de .70. No Brasil, os fatores “Emoção (aluno)”, “Irrelevante (ruim)” e “Irrelevante (ignorar)” apresentaram valor inferior a .70. Já na Nova Zelândia, isso aconteceu com os fatores “Irrelevante (ignorar)” e “Responsabilização (escola)”. Dessa forma, a maioria dos fatores teve uma estimação de confiabilidade interna suficiente para fins de comparação entre as duas amostras. Entretanto, vale destacar que os coeficientes de consistência interna possuem a limitação de considerar que todos os itens medem igualmente o construto. Além disso, a análise da consistência interna não pode ser considerada como prova cabal da validade de construto do teste (PASQUALI, 2003). Apontamos ainda que o fato de “Emoção (aluno)” e “Responsabilização (escola)” possuírem apenas dois itens cada pode ter influenciado os valores dos seus respectivos coeficientes.

Antes de discutir as evidências de validade do instrumento, lembramos o modelo completo proposto por Brown *et al.* (2009), que possui quatro fatores de segunda ordem (Melhora, Emoção, Irrelevante e Responsabilização) e oito fatores de primeira ordem (Melhora-aluno, Melhora-professor, Emoção-aluno, Emoção-sala, Irrelevante-ruim, Irrelevante-ignorar, Responsabilização-aluno, Responsabilização-escola). Assim, esperávamos encontrar evidência convergente entre os fatores I e II, III e IV, V e VI, VII e VIII, pois todos eles compartilham um mesmo fator de segunda ordem (ex: “Emoção-aluno” e “Emoção-sala”). Já entre os fatores que não compartilham um mesmo fator de segunda ordem, esperávamos encontrar evidência discriminante (ex: “Melhora-aluno” e “Irrelevante-ruim”). Nesse sentido, apresentamos na TAB. 6 a matriz de correlação completa entre os fatores do SCoA brasileiro.

TABELA 6
Correlação entre os fatores do SCoA Brasileiro

Fatores do SCoA	I	II	III	IV	V	VI	VII	VIII
I. Melhora (aluno)	1.00							
II. Melhora (professor)	.62	1.00						
III. Emoção (aluno)	.60	.87	1.00					
IV. Emoção (sala)	.39	.51	.63	1.00				
V. Irrelevante (ruim)	-.47	-.61	-.68	-.29	1.00			
VI. Irrelevante (ignorar)	-.63	-.47	-.50	-.21	.79	1.00		
VII. Respons (aluno)	.58	.81	.83	.54	-.53	-.44	1.00	
VIII. Respons (escola)	.50	.72	.62	.42	-.44	-.35	.69	1.00

Nota: correlações estimadas com Análise Fatorial Confirmatória.

Quanto à evidência convergente, os valores foram de acordo com o esperado, apresentando correlações médias e fortes. As correlações entre os fatores I e II, III e IV, V e VI, VII e VIII variaram entre .62 e .79. A correlação maior foi entre os fatores “Irrelevante-ruim” e “Irrelevante-ignorar” ($r = .79$).

Já com relação à evidência discriminante, os valores não foram de acordo com o esperado, pois algumas correlações foram altas. São alguns exemplos: “Emoção (aluno)” e “Melhora (professor)” ($r = .87$); “Melhora (professor)” e “Responsabilização (aluno)” ($r = .81$). As correlações entre os fatores que não compartilham um mesmo fator de segunda ordem variaram entre -.21 e .87. Tomados em conjunto, esses resultados indicam que alguns fatores são redundantes. Também chama a atenção o fato de que algumas correlações tiveram valores superiores aos encontrados na evidência convergente. No entanto, como já destacamos na seção 5.2.3, existem métodos mais robustos para a validação de um questionário como, por exemplo, a Análise Fatorial

Confirmatória (AFC). Nesse sentido, a próxima seção descreve de forma detalhada os resultados da AFC.

6.3.2 - Análise Fatorial Confirmatória

Nessa pesquisa também utilizamos a Análise Fatorial Confirmatória (AFC), por ser um procedimento de estimação mais robusta para a validação do questionário.

Com relação às concepções de avaliação dos alunos (itens 1 ao 32), analisamos os índices de ajuste de três modelos: quatro fatores (ou seja, agrupamos os oito fatores); oito fatores; e o modelo completo proposto por Brown *et al.* (2009) com quatro fatores de segunda ordem (Melhora, Emoção, Irrelevante e Responsabilização) e oito fatores de primeira ordem (Melhora-aluno, Melhora-professor, Emoção-aluno, Emoção-sala, Irrelevante-ruim, Irrelevante-ignorar, Responsabilização-aluno, Responsabilização-escola). A solução com fatores de segunda ordem foi inadmissível na Análise Fatorial Confirmatória.

O *software* MPLUS versão 3.11 foi utilizado para realizar a AFC do SCoA. O método de Máxima Verossimilhança, que é o estimador padrão do MPLUS, assume uma normalidade multivariada. Esse pressuposto é violado com o SCoA, que contém itens do tipo Likert de seis pontos. Assim, a AFC foi conduzida usando um estimador robusto do MPLUS para a análise de indicadores categóricos denominado *Mean and variance-adjusted weighted least squares estimator* (WLSMV). O WLSMV fornece uma estimação sem viés (*unbiased estimates*) para escalas Likert de diferentes números de categorias e graus de simetria e curtose (FLORA; CURRAN, 2004).

A TAB. 7 indica os valores encontrados para as respostas ao SCoA brasileiro.

TABELA 7
Comparação de índices de ajuste dos modelos da AFC: Concepções de avaliação

Modelos	χ^2	gl	CFI	TLI	Gamma _{hat}	RMSEA	SRMR
4 fatores	2198.634	127	.64	.89	.85	.147	.087
8 fatores	1116.312	138	.83	.95	.92	.097	.057

Nota: Todos os valores χ^2 foram estatisticamente significantes ($p < .0000$).

Como apontado na TAB.7, o modelo de quatro fatores teve um ajuste muito ruim e o modelo com oito fatores teve o melhor ajuste. 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). Esses índices demonstram o ajuste global do modelo (em que medida a hipótese teórica se ajusta aos dados observados). Juntos esses índices sugerem que as respostas ao SCoA brasileiro se ajustam bem à estrutura fatorial proposta de oito fatores. As cargas fatoriais padronizadas de todos os itens relacionados às concepções de avaliação do SCoA brasileiro (itens 1 ao 32) são apresentadas na TAB. 8.

TABELA 8
Cargas fatoriais padronizadas dos itens do SCoA Brasileiro: Concepções de avaliação

<i>Melhora (aluno)</i>		<i>Melhora (prof)</i>		<i>Emoção (aluno)</i>		<i>Emoção (sala)</i>	
Item	Carga	Item	Carga	Item	Carga	Item	Carga
1	.64	5	.75	6	.66	2	.82
10	.69	8	.86	31	.84	12	.84
14	.72	9	.86			17	.70
15	.86	23	.70			21	.68
19	.78	27	.51			25	.80
		30	.70			28	.78
<i>Irrelev (ruim)</i>		<i>Irrelev (Ignorar)</i>		<i>Respons (aluno)</i>		<i>Respons (escola)</i>	
Item	Carga	Item	Carga	Item	Carga	Item	Carga
3	.60	7	.57	4	.65	11	.74
13	-.18	29	.62	16	.75	24	.88
18	.40	32	.77	20	.80		
22	.33						
26	.87						

Como indicado na TAB. 8, a maioria dos itens teve cargas fatoriais altas. O item 13 foi uma exceção. Esse item se mostrou problemático em todas as análises realizadas. Na AFC, por exemplo, ele apresentou uma variância residual de .97. Dessa forma, esse item precisa ser modificado ou excluído em futuras aplicações do questionário. A hipótese que temos sobre a causa do mau funcionamento do item 13 na versão brasileira se refere à tradução. O item original em inglês é “*Assessment interferes with my learning*” e a tradução para o português foi “A avaliação interfere no meu aprendizado”. Acreditamos que a palavra “interfere” pode apresentar uma ambiguidade, significando

tanto algo como “faz diferença, modifica o meu aprendizado” ou “atrapalha meu aprendizado”. Talvez isso tenha confundido os respondentes do questionário.

Por fim, uma mesma crítica que foi feita na análise das definições de avaliação pode ser feita aqui. Dois fatores “Emoção (aluno)” e “Responsabilização (escola)” possuem somente dois itens. Tecnicamente, não é muito adequado ter um fator só com dois itens. Por isso, recomendamos que novos itens sejam incluídos nesses fatores em futuras análises.

6.3.3 - Análise Fatorial Exploratória

Como já discutimos anteriormente, encontrar um bom ajuste do modelo não significa provar empiricamente que o modelo está “correto”. Vários modelos podem se ajustar aos dados. Assim, consideramos prudente realizar uma Análise Fatorial Exploratória (AFE) e verificar se seu resultado seria congruente com o modelo testado na Análise Fatorial Confirmatória (AFC). Apesar de estarmos fazendo uso de teoria e pesquisas anteriores, avaliamos que esse procedimento foi necessário devido às possíveis diferenças entre Brasil e Nova Zelândia. Além disso, essa é uma primeira pesquisa com dados brasileiros. Esse procedimento tem como objetivo verificar se algum modelo alternativo surge dos dados empíricos.

O *software* MPLUS versão 5.1 foi utilizado para realizar a AFE do SCoA, que foi conduzida usando um estimador robusto do MPLUS para a análise de indicadores categóricos (*Mean and variance-adjusted weighted least squares estimator* - WLSMV). Usamos um método de rotação oblíqua denominado *OBLIMIN*. Segundo Hair *et al.* (2005), a rotação fatorial oblíqua é computada de forma que os fatores extraídos são correlacionados. Já na rotação fatorial ortogonal cada fator é independente (ortogonal) em relação a todos os outros (a correlação entre eles é assumida como sendo zero). Por isso realizamos uma rotação fatorial oblíqua (os fatores do SCoA estão correlacionados). Para Field (2005), o uso de rotação ortogonal com dados de Ciências Humanas não parece ter sentido algum. Não existem, por exemplo, variáveis educacionais ou psicológicas que não sejam correlacionadas. Assim, alguns autores argumentam que esse tipo de rotação fatorial nunca deveria ser utilizado em Ciências Humanas. O pesquisador precisaria ter evidências teóricas ou empíricas muito fortes de que os fatores não são correlacionados. Essa é uma falha metodológica encontrada com frequência em pesquisas que fazem uso de análise fatorial.

O MPLUS apresenta os resultados da AFE de uma forma um pouco diferente de outros programas como o SPSS. Na AFE são fornecidos os mesmos índices de ajuste da

AFC, com o intuito de comparar soluções com número de fatores diferentes. O mesmo critério de corte da AFC pode ser utilizado para decidir o número de fatores. Assim, o pesquisador precisa estabelecer um número mínimo e máximo da quantidade de fatores que devem ser extraídos. A TAB. 9 mostra resultados de oito soluções diferentes de AFE.

TABELA 9
Comparação de índices de ajuste da AFE: Concepções de avaliação

Modelos	χ^2	gl	CFI	TLI	Gamma _{hat}	RMSEA	SRMR
1 fator	3375.351	119	.44	.82	.79	.190	.109
2 fatores	2166.646	140	.65	.90	.86	.138	.073
3 fatores	1567.973	148	.76	.94	.89	.113	.054
4 fatores	1169.914	148	.82	.95	.92	.096	.041
5 fatores	890.288	149	.87	.97	.94	.081	.034
6 fatores	672.023	143	.91	.97	.96	.070	.030
7 fatores	527.681	137	.93	.98	.97	.061	.025
8 fatores	440.517	135	.95	.98	.97	.055	.022

Nota: Todos os valores χ^2 foram estatisticamente significantes ($p < .0000$).

Como apontado na TAB. 9, as análises com 6, 7 e 8 fatores tiveram todos os cinco índices indicando um bom ajuste do modelo. Dentre os três, o modelo com 8 fatores foi o que apresentou melhor ajuste aos dados. Apesar do MPLUS oferecer esses índices de ajuste, destacamos que esses resultados se referem a uma AFE. Vale lembrar que nesse tipo de análise fatorial o pesquisador deixa os dados observados determinarem o modelo fatorial subjacente a posteriori. Numa técnica exploratória deixamos “os dados falarem por si só”. Nós não pré-determinamos uma estrutura. Assim, apesar da nossa revisão da literatura e da AFC nos levarem a esperar que os construtos latentes do SCoA existem no Brasil, nós permitimos via AFE uma junção puramente empírica dos dados para nos certificar disso. Nesse tipo de análise, o MPLUS oferece esse recurso bastante interessante de comparar várias análises fatoriais exploratórias diferentes. A TAB. 10 apresenta as cargas fatoriais de todos os itens relacionados às concepções de avaliação do SCoA brasileiro no modelo com 8 fatores, que apresentou o melhor ajuste.

TABELA 10
AFE: cargas fatoriais dos itens do SCoA Brasileiro - Concepções de avaliação

Item	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5	Fator 6	Fator 7	Fator 8
1	-.014	-.006	.667	.044	-.047	-.032	-.003	.027
2	.076	.832	.010	.001	.013	.020	-.046	-.055
3	-.149	.072	.077	-.038	.417	-.261	.213	-.090
4	.318	.154	-.028	.176	.141	.236	.118	-.022
5	.509	.091	-.018	.129	-.109	-.021	.049	.200
6	.239	.203	.017	-.085	-.079	.329	-.023	.167
7	-.050	.033	-.286	.073	.457	.077	.051	-.003
8	.947	-.018	-.007	-.009	-.006	.023	.006	-.002
9	.778	.004	.102	.083	-.030	-.030	.018	.007
10	.034	.028	.597	.078	-.085	-.021	-.074	.077
11	.080	.063	.035	.827	.008	-.081	-.085	-.039
12	.061	.868	.053	.067	.018	.016	-.099	-.123
13	.061	-.023	.135	.020	.026	-.008	.214	-.099
14	.043	.024	.801	-.010	.015	.015	-.053	-.039
15	.043	.059	.664	.021	-.012	.217	.057	-.046
16	.221	.055	.028	.146	.022	.239	.407	.045
17	.022	.685	.009	-.077	.028	-.009	.159	.102
18	.031	.019	.147	-.041	.452	-.191	.271	-.193
19	.013	-.028	.688	-.004	-.059	.006	.089	.168
20	.267	.056	-.002	.048	-.224	.145	.462	.064
21	.058	.409	-.068	-.085	-.101	.148	.090	.299
22	-.076	.043	.003	-.011	.153	-.305	.173	.080
23	.065	.052	.129	.126	-.042	.181	.052	.474
24	-.036	-.028	-.025	.711	-.054	.150	.137	.121
25	-.132	.868	-.032	.039	-.047	-.033	.069	.092
26	-.139	-.065	-.017	-.024	.692	-.077	.031	-.018
27	.104	.010	.125	.112	.110	-.048	-.002	.639
28	.009	.449	.102	.036	-.025	.092	.007	.331
29	.027	-.055	-.224	-.022	.600	.142	-.117	.149
30	.003	-.054	.183	.175	-.006	.558	.107	.026
31	.110	.116	.145	.040	-.068	.620	.055	.038
32	-.005	.032	-.073	-.095	.615	-.029	-.164	.057

Nota: as cargas fatoriais mais altas estão em negrito. Método de rotação: *OBLIMIN*.

Os resultados da TAB. 10 foram próximos aos resultados da estrutura fatorial proposta no SCoA da Nova Zelândia. Dessa forma, essa é mais uma evidência de que essa estrutura fatorial se mantém com os dados brasileiros obtidos com estudantes do ensino superior. A estrutura fatorial original do SCoA é:

- Melhora (aluno) 1, 10, 14, 15, 19.
- Melhora (professor) 5, 8, 9, 23, 27, 30.
- Emoção (aluno) 6, 31.
- Emoção (sala) 2, 12, 17, 21, 25, 28.
- Irrelevante (ruim) 3, 13, 18, 22, 26.

- Irrelevante (ignorar) 7, 29, 32.
- Responsabilização (aluno) 4, 16, 20.
- Responsabilização (escola) 11, 24.

Os resultados da AFE foram semelhantes:

- Fator 1: itens 4, 5, 8, 9. (3 itens de Melhora-professor e 1 de Responsabilização-aluno).
- Fator 2: itens 2, 12, 17, 21, 25, 28. (igual a Emoção-sala)
- Fator 3: itens 1, 10, 14, 15, 19. (igual a Melhora-aluno)
- Fator 4: itens 11, 24. (igual a Responsabilização-escola)
- Fator 5: itens 3, 7, 18, 26, 29, 32. (itens de Irrelevante-ruim e Irrelevante-ignorar)
- Fator 6: itens 6, 22, 30, 31. (itens de Emoção-aluno, Melhora-professor e Irrelevante-ruim)
- Fator 7: itens 13, 16, 20. (quase igual a Responsabilização-aluno. Lembrando que o item 13 teve carga fatorial baixa em todos os fatores)
- Fator 8: itens 23, 27. (itens de Melhora-professor)

Portanto, concluímos que nenhum modelo alternativo surgiu da empiria pura dos dados. Lembrando que os índices de ajuste da AFC foram muito bons, também concluímos que o nível de ensino, a língua e o país não interferiram na forma como os itens se agruparam em variáveis latentes (os fatores do SCoA). Os mesmos construtos latentes existem em duas culturas bem diferentes (Brasil e Nova Zelândia) e isso fornece uma evidência de validade adicional para o SCoA.

6.3.4 - Análises adicionais sobre as concepções de avaliação

Para uma compreensão adicional das concepções de avaliação, as médias das respostas para cada fator foram examinadas (FIG. 5). 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).

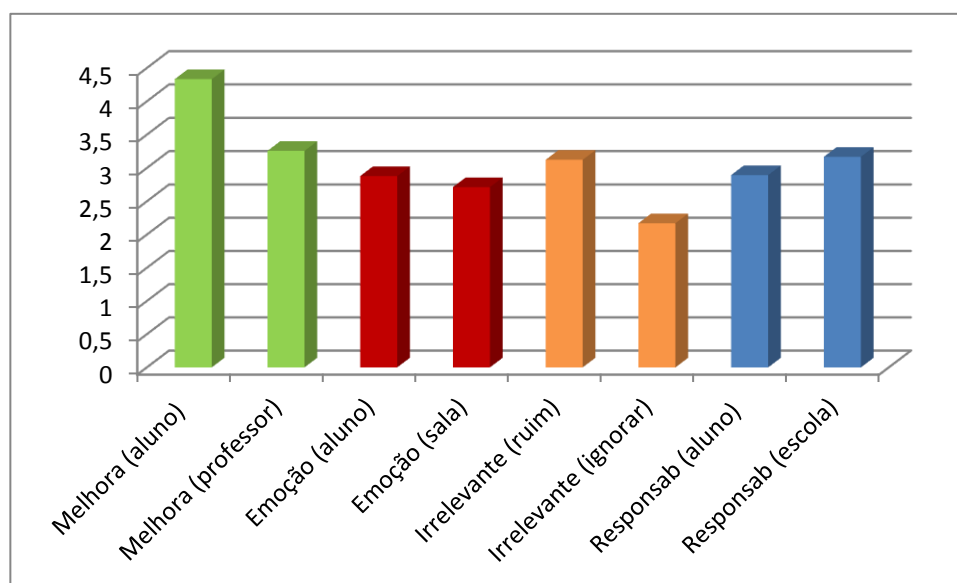


FIGURA 5 - Médias das respostas das concepções de avaliação

Nota: N=756

Como a FIG. 5 aponta, os alunos parecem conceber que a avaliação primeiramente 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.

Em outras palavras, tomando como referência todos os alunos da amostra, eles 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”*.

6.3.5 - Diferenças entre a IES pública e a IES privada

Com o objetivo de comparar as concepções de avaliação de alunos pertencentes a uma IES pública e uma IES privada (uma Universidade Federal e um Centro Universitário privado, ambas IES do estado de Minas Gerais) (hipótese 5), realizamos uma Análise de Variância (ANOVA). A TAB. 11 compara as médias das oito concepções de avaliação.

TABELA 11
Diferenças entre IES pública e IES privada nos fatores do SCoA

Fatores	IES	Média	DP
Melhora (aluno)	Pública	4.10	1.09
	Privada	4.48	1.00
Melhora (professor)	Pública	2.75	.86
	Privada	3.58	1.05
Emoção (aluno)	Pública	2.42	1.09
	Privada	3.17	1.30
Emoção (sala)	Pública	2.55	1.01
	Privada	2.82	1.10
Irrelevante (ruim)	Pública	3.39	.73
	Privada	2.95	.86
Irrelevante (ignorar)	Pública	2.19	.92
	Privada	2.15	.99
Responsabilização (aluno)	Pública	2.37	.99
	Privada	3.23	1.28
Responsabilização (escola)	Pública	2.77	1.16
	Privada	3.42	1.35

Uma das pressuposições (*assumptions*) da ANOVA é a homogeneidade de variância. Essa pressuposição foi violada em todas as comparações, com a exceção da escala “Melhora (aluno)”, $F(1, 754) = 24.96$, $p < .000$. A ANOVA se mostra um procedimento resistente a essa violação quando os tamanhos das amostras são iguais (FIELD, 2005). Como isso não aconteceu nos nossos dados (Pública, $N= 297$, Privada, $N= 459$), realizamos as comparações entre os tipos de IES por meio de um teste robusto de igualdade de médias: o valor de F de Welch. A TAB. 12 mostra a comparação das médias de sete escalas do SCoA brasileiro.

TABELA 12
Teste robusto de igualdade de médias: tipo de IES

Fatores	Welch F	gl 1	gl 2	p
Melhora (professor)	140.551	1	711.844	.000
Emoção (aluno)	74.254	1	706.007	.000
Emoção (sala)	12.064	1	670.972	.001
Irrelevante (ruim)	55.773	1	703.227	.000
Irrelevante (ignorar)	.298	1	665.735	.585
Responsabilização (aluno)	107.248	1	729.922	.000
Responsabilização (escola)	49.590	1	694.602	.000

A análise das TAB. 11 e 12 indica que a única diferença que não foi estatisticamente significativa foi a escala “Irrelevante (ignorar)”. Os alunos do Centro Universitário privado acreditam mais que a avaliação melhora o ensino-aprendizagem e torna alunos e escolas responsáveis do que os alunos da universidade pública (Melhora-aluno; Melhora-professor; Responsabilização-aluno; Responsabilização-escola). Além disso, os alunos da instituição pública acreditam em um maior grau que a avaliação é algo ruim ou injusto (Irrelevante-ruim). Por fim, os alunos do Centro Universitário privado percebem uma maior associação da avaliação a benefícios como a motivação dos colegas e o clima da turma (Emoção-sala). Os alunos da instituição privada também concebem a avaliação como uma experiência mais agradável quando comparados a alunos da instituição pública (Emoção-aluno).

Considerando nossa hipótese 5, não esperávamos encontrar diferenças nas concepções de avaliação dos alunos quando comparamos instituições de ensino públicas e privadas. Não encontramos elementos na nossa revisão da literatura que nos apontassem possíveis diferenças (essa temática não foi abordada em nenhum trabalho que pesquisamos). Existem algumas explicações possíveis para tais resultados.

Em primeiro lugar, a literatura descreve de forma bastante crítica a avaliação no ensino superior do Brasil, destacando tanto um caráter político quanto uma vinculação com o poder (ver capítulo 1). Destacamos que quase sempre essa visão crítica sobre a avaliação é oriunda de pesquisadores de IES públicas. Assim, é possível que os professores dessas instituições transmitam essa visão para os estudantes.

Em segundo lugar, 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, considerando a diferença de resultados no fator “Emoção”, lembramos uma questão abordada na seção 3.3 dessa pesquisa: embora pudéssemos esperar que um impacto emocionalmente positivo nos alunos (individualmente ou na sala de aula) poderia elevar 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 auto-regulada, Brown, Peterson e Irving (2009) argumentam que somente a concepção de melhora parece estar alinhada com uma resposta positiva de auto-regulação do aprendizado. Dessa forma, apesar do fato de que os alunos do Centro Universitário privado acreditam mais do que os alunos da Universidade Federal na concepção “Melhora”, eles também acreditam mais nas concepções “Emoção” e “Responsabilização” (que seriam potencialmente concepções com uma relação negativa com a auto-regulação do aprendizado). Portanto, consideramos que todas essas relações tomadas em conjunto possam oferecer explicações plausíveis sobre as diferenças entre as duas IES.

6.4 - Resultados qualitativos

Após os alunos responderem a questão “Quando você pensa na palavra AVALIAÇÃO, quais tipos de atividades avaliativas vêm a sua mente ?” marcando as opções que quisessem, havia a pergunta aberta “Algo mais: Escreva com letra legível”. Nesse local, os estudantes tinham liberdade para incluir atividades que não tinham sido mencionadas na lista. Apresentamos a seguir um apanhado geral dessas respostas. Nossos comentários estão entre parênteses e em itálico.

- Quando distribui pontos com base na participação em aula.
- O professor aplica prova individual e sem consulta (*na nossa lista existe uma opção que diz “prova em dupla”. Nas outras opções talvez fosse prudente mencionar explicitamente que a prova é individual*).
- Apresentação oral de trabalho (*sugestão dada por vários alunos*)
- O professor avalia meu comportamento, minha conduta.

- Avaliação de um trabalho interdisciplinar.
- Grupos diferentes, estudar como pesquisa, ir a lugares. *(essa sugestão parece indicar a avaliação de visitas técnicas, uma prática realizada em muitos cursos de graduação).*
- O professor avalia o aluno durante uma aula prática ou jogo.
- Apresentação de seminário *(sugestão dada por vários alunos)*
- Testes Oraís - ex. língua estrangeira *(essa resposta serve para mostrar que às vezes um questionário com práticas avaliativas mais universais não consegue captar especificidades de alguns cursos. Essa atividade se refere ao curso de letras)*
- Resenhas
- Trabalhos Críticos
- Dinâmicas
- Debates ou discussões.
- Provas práticas - ex. anatomia, neuroanatomia *(essa resposta serve para mostrar que às vezes um questionário com práticas avaliativas mais universais não consegue captar especificidades de alguns cursos. Essa atividade se refere a cursos da área de biológicas)*
- Relatórios de atividades práticas *(tipo de atividade presente em vários cursos)*
- Filmes e discussão *(prática que nos parece cada vez mais adotada em diversos cursos. Quanto à avaliação e pontuação dessa atividade, não temos certeza se é comum).*
- O professor deixa os alunos discutirem em grupo e atua como mediador, avaliando as discussões.

Consideramos que os alunos forneceram valiosas sugestões para pesquisas futuras com o SCoA. Assim, novos tipos de atividades avaliativas podem ser incluídas e testadas na parte complementar do questionário. São alguns exemplos: apresentação oral de trabalho, apresentação de seminário e relatórios de atividades práticas. Atividades muito específicas como testes orais de uma língua estrangeira também propiciam uma reflexão para futuras versões do instrumento: a possibilidade de apresentação aos estudantes de atividades mais recortadas de acordo com cada área do conhecimento. Por fim, as respostas corroboraram nossa hipótese de que o portfólio ainda não é uma prática muito utilizada no ensino superior, pois nenhum estudante mencionou tal atividade avaliativa.

6.4.1 - Análise de conteúdo

O final do SCoA possui a seguinte questão: “Você tem algum comentário ?”. Nessa questão aberta, os alunos tinham liberdade para mencionar quaisquer aspectos que considerassem relevantes.

A partir da análise das respostas, identificamos tendências predominantes no discurso dos estudantes. Isso possibilitou a elaboração das seguintes categorias analíticas: a) O papel da avaliação: percepções dos alunos sobre as funções da avaliação; b) Percepções sobre os efeitos da avaliação; c) Percepções sobre a Educação: concepções implícitas nos discursos analisados. Os trechos a seguir serão interpretados a partir dessas categorias.

O papel da avaliação: percepções dos alunos sobre as funções da avaliação

Como já discutimos na seção 1.1, Jorba e Sanmartí (2003) apresentam duas funções da avaliação: uma de caráter social, de seleção e classificação; outra de caráter pedagógico, de ajuste do processo ensino-aprendizagem, de reconhecimento de mudanças que podem ser introduzidas progressivamente num determinado processo. Algumas respostas dos estudantes indicam uma ênfase na função de seleção e classificação:

A nota de prova, ou do final da etapa (histórico), não vão me ajudar no mercado de trabalho, que é o objetivo de se fazer um curso (trabalhar). Estas notas, meramente vão me aprovar de uma etapa para outra, sem dizer na verdade qual o meu nível de conhecimento.

Prova não prova nada e aluno que tirou nota baixa ou foi reprovado nem sempre é malandro.

Essas falas dos alunos também parecem ligadas à concepção de responsabilização (a avaliação para aprovar ou reter). No entanto, alguns alunos mencionam a função pedagógica da avaliação (sua dimensão formativa):

A avaliação não deveria ter o único objetivo de atribuir nota aos alunos, mas também de dar a oportunidade ao aluno dele saber o que errou e porque errou, de modo que o mesmo possa aprofundar no que ainda lhe proporciona dificuldades.

A avaliação deveria ser realizada para apontar as lacunas que ficaram no aprendizado, e não como forma sádica de controle como alguns professores a utilizam. Deveria ser utilizada para entender o que você

aprendeu e fornecer dados para ampliar e não somente testar e descobrir o que você não sabe e pronto.

O professor dificilmente dá um trabalho no qual gaste tempo dando sugestões e tirando dúvidas, e não há devolutivas em relação ao que o aluno precisa melhorar e do que ele assimilou, na verdade, na maioria das vezes recebemos só a nota, não temos acesso a prova para sabermos nossos erros e acertos. Ter avaliação é fundamental, mas não dessa forma.

Muitas vezes, nas faculdades públicas, os alunos não têm retorno das provas e trabalhos feitos, por isso a minha desilusão com relação à avaliação. Mas penso que se fosse mais bem planejada e houvesse correção e discussão junto aos alunos, seria mais proveitoso.

As respostas das questões das avaliações deveriam ser conferidas e discutidas com o aluno para verificar se ele realmente aprendeu. Desabafo: Aqui no (prédio de Ciências Biológicas da Universidade Federal) tem professores que não entregam a prova e não comentam sobre ela. Isso é muito comum e é um absurdo. Obs: A minha turma sempre se ajuda e estuda em grupo!

Assim, o que mais chama a atenção nesses trechos é o distanciamento entre como “as coisas deveriam ser” e como elas são na realidade. Ou seja, os alunos têm consciência da função pedagógica da avaliação e reconhecem que ela não acontece: “A avaliação deveria ser realizada para apontar as lacunas que ficaram no aprendizado”, deveria “dar a oportunidade ao aluno dele saber o que errou e porque errou”. Além disso, até mesmo aspectos básicos da avaliação formativa parecem ser negligenciados por alguns docentes, como a questão do *feedback*: “não há devolutivas em relação ao que o aluno precisa melhorar”, “professores que não entregam a prova e não comentam sobre ela”. Nesse sentido, pesquisas centradas nos discentes têm mostrado que eles cada vez mais demandam de seus professores um *feedback* cognitivo/acadêmico.

Percepções sobre os efeitos da avaliação

Um dos pontos que se destacou no discurso dos estudantes foi uma forte reação emocional negativa:

Muitas vezes o nervosismo atrapalha e faz confusão em nossa cabeça.

Quando muitas vezes sabemos o conteúdo e por pressão, ou medo, enfim, fatores psicológicos não fazemos boa prova.

A palavra avaliação traz pânico, nervosismo, desespero, medo tristeza e a prova.

Sobre pressão não há devida concentração.

Eu fico com medo.

Assim, reações emocionais como “pânico” e “medo” apareceram com frequência. No entanto, mais uma vez ficou evidente que uma gama extensa de concepções de avaliação foi identificada nas experiências e pensamentos dos alunos. Os trechos a seguir ilustram essa questão:

Avaliações "em grupo" são injustas, pois, nem todos trabalham.

A forma de avaliação do (Centro Universitário privado) é muito complicada para a maioria das pessoas, que trabalham e não dispõem, de tempo para estudar e fazer trabalhos complicados. A avaliação deveria ser durante as aulas (Trabalhos em sala, exercícios...)

Como o nosso curso é no turno da noite, muitos alunos trabalham e o tempo se torna escasso para cumprir as tarefas. Gostaria que este ponto fosse avaliado para uma melhoria em nosso curso.

Embora eu não me sinta bem durante uma avaliação e apesar dela não ter valor significativo quanto a inteligência ou capacidade. Eu reconheço que é uma das formas de avaliar a qualidade do aprendizado...

Mas eu aprecio muito a autoavaliação e a avaliação feita pelo professor conjuntamente comigo sobre o meu trabalho.

Assim, os três primeiros trechos se referem a uma concepção de irrelevância. A avaliação é considerada injusta por dois motivos: primeiro porque em trabalhos em grupo “nem todos trabalham”; segundo, porque nem sempre leva em consideração realidades complexas típicas do contexto brasileiro como a dos estudantes trabalhadores: “muitos alunos trabalham e o tempo se torna escasso”. O penúltimo trecho revela concepções de irrelevância e responsabilização. Por fim, o último remete a um impacto emocionalmente positivo no aluno, a um prazer pessoal no ato de ser avaliado.

Percepções sobre a Educação: concepções implícitas nos discursos analisados

Algumas respostas dos alunos indicam que eles têm consciência de que a forma como um professor avalia é reflexo de sua prática pedagógica como um todo. Ou seja, a maneira de avaliar está diretamente ligada à uma concepção de educação.

É inútil algumas avaliações aplicadas pelo professor, decorebas sem sentido, que só faz com que nós alunos nos sintamos inúteis.

A avaliação da forma como ela é dada atualmente não mede o conhecimento de toda a turma, e não valoriza o pensamento crítico, uma vez que para tirarmos boas notas devemos reproduzir um determinado pensamento ou autor eleito pelo professor.

Infelizmente a maioria dos professores acham que avaliar é aplicar uma prova, desconsiderando todo o processo de aprendizagem e até mesmo o esforço do aluno.

O professor aplica prova e nem sequer reflete se ele realmente ensinou algo.

Professor só porque tem estudos e sabe muito nem sempre sabe ensinar e conseguir com que os alunos entendam e aprendam. Alunos de faculdade não sabem tudo, se não não estariam numa faculdade estudando.

Consideramos que por trás das ações de qualquer professor, há sempre um conjunto de ideias que as orienta (tenha o professor consciência disso ou não). A avaliação não é um fenômeno isolado na prática do professor (WEISZ, 1999). Assim, um professor que faz avaliações do tipo “decorebas” ou que exige do aluno “reproduzir um determinado pensamento ou autor eleito pelo professor” provavelmente reproduz essa atitude pedagógica tradicional em seu cotidiano. As afirmações de que o professor “nem sequer reflete se ele realmente ensinou algo” e “nem sempre sabe ensinar e conseguir com que os alunos entendam e aprendam” nos remetem para outro ponto relevante: o fato de um profissional ter um grande conhecimento sobre determinada área não faz com que ele automaticamente seja um bom professor. O domínio do processo pedagógico é de outra ordem (precisa ser um processo deliberado e sistemático) (SAVIANI, 1984). Ou seja, é necessária uma tomada de consciência por parte do educador desse processo. Daí a importância do trabalho educativo com o devido embasamento teórico (inclusive a elaboração de avaliações). Nesse sentido, diversas pesquisas apontam que os professores não estão suficientemente familiarizados com as teorias de cunho pedagógico e possuem deficiências na sua formação (VEIGA; RESENDE; FONSECA, 2000).

Ainda considerando a relação entre prática pedagógica e avaliação, algumas respostas evidenciaram uma forte concepção de responsabilização do professor:

Eu reconheço que é uma das formas de avaliar a qualidade do aprendizado e a qualidade dos futuros profissionais e até mesmo da qualidade e capacidade de um professor.

Acredito que os professores deveriam ser abertos a avaliação também, mas não tem como acontecer aqui no (Centro Universitário privado), assim como somos avaliados no nosso desempenho, eles

também deveriam ser, porque acho que o ensino se torna muito mais fácil quando eles fazem o que quiser sem levar em conta as falhas de seu ensino.

Dessa forma, a avaliação pode servir para atestar a “qualidade e capacidade de um professor”. Mais uma vez aparece aqui a ideia implícita de que a forma como um professor avalia é reflexo de sua prática pedagógica. Portanto, assim como os estudantes são avaliados, os professores “também deveriam ser” (esse último trecho também revela outro aspecto interessante: o imaginário do aluno de que os professores não são avaliados).

6.5 - Relação entre as concepções e as definições de avaliação: Modelagem de Equações Estruturais

No presente trabalho também analisamos como as concepções de avaliação predizem as definições de avaliação dos alunos (hipótese 4). Uma combinação de procedimentos exploratórios e confirmatórios foi utilizada. No primeiro momento, foi realizado um Escalonamento Multidimensional exploratório para descobrir a estrutura subjacente às definições de avaliação dos alunos. Esse procedimento exploratório foi necessário porque cinco dos 12 itens originais do questionário foram modificados para se adequarem à realidade de alunos do ensino superior brasileiro. O modelo de medida encontrado foi testado posteriormente em uma análise fatorial confirmatória. Com relação às concepções de avaliação, somente um item foi excluído do original do SCoA (item 33).

Para testar um modelo estrutural, utilizamos inicialmente o modelo encontrado por Brown *et al.* (2009) (práticas formais e interativo-informais), pois o critério para a seleção de variáveis deve ser sempre teórico e baseado em pesquisas anteriores. No entanto, a solução para esse modelo foi inadmissível nos dados da amostra brasileira.

Assim, foi iniciado um novo procedimento exploratório por meio da utilização de regressões múltiplas. Cada uma das três definições de avaliação foi utilizada como variável dependente em um modelo de regressão múltipla onde as oito concepções de avaliação eram as variáveis independentes. Foram selecionados coeficientes de regressão que eram significantes para serem adicionados na Modelagem de Equações Estruturais. Vários modelos foram testados.

O modelo apresentado na FIG. 6 foi o que apresentou o melhor ajuste: χ^2 (205, $N= 756$) = 940.095, $p < .0000$, CFI= .90, TLI= .96, Gamma hat= .96, RMSEA=.069,

SRMR=.062.²¹ Cinco índices indicam um bom ajuste, sugerindo que as relações estruturais propostas se ajustam bem ao modelo. O *software* MPLUS versão 3.11 foi utilizado para realizar a Modelagem de Equações Estruturais. A análise foi conduzida usando o estimador robusto do MPLUS para a análise de indicadores categóricos chamado *Mean and variance-adjusted weighted least squares estimator* (WLSMV).

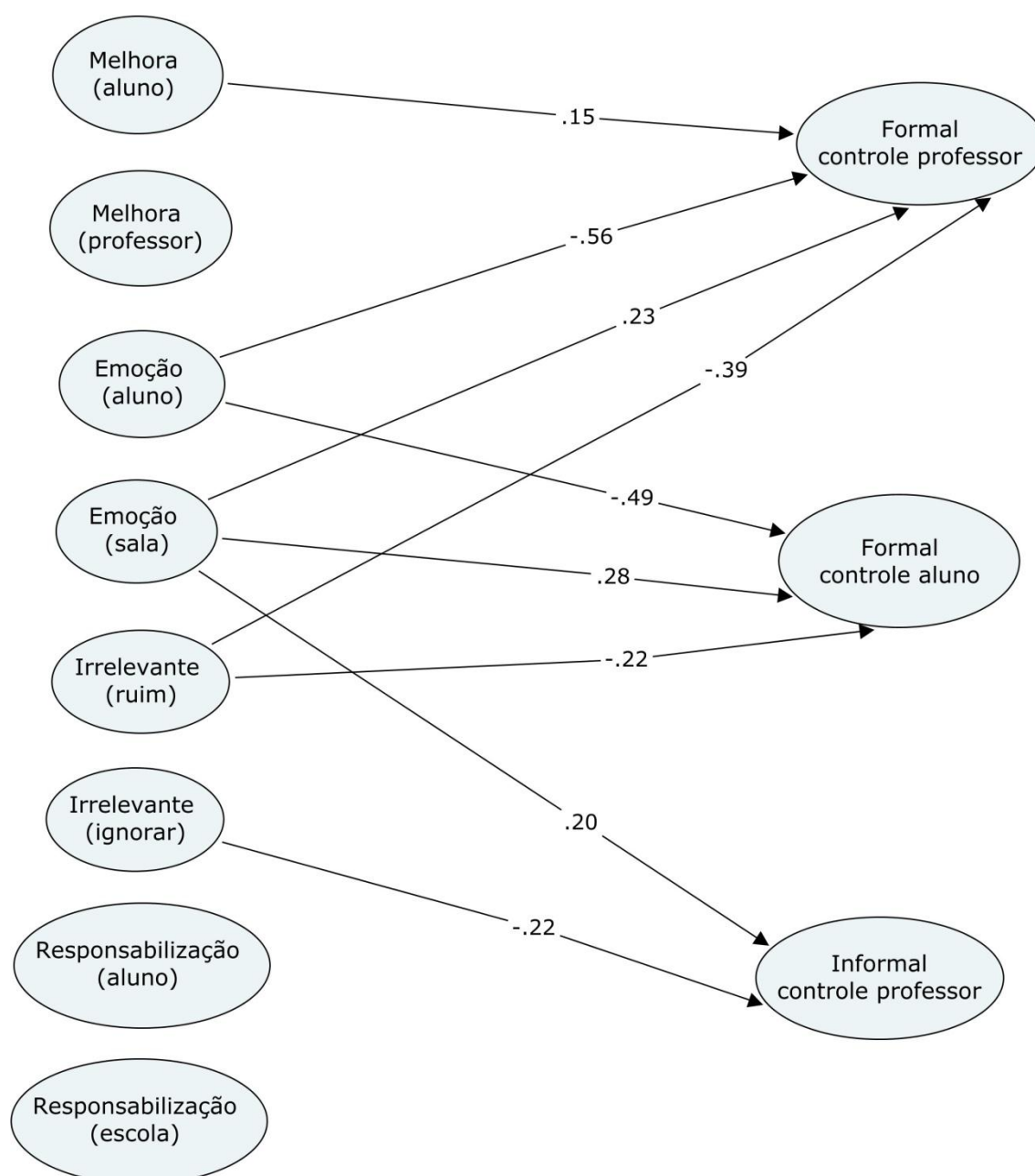


FIGURA 6 - Modelagem de Equações Estruturais

Nota: Todos os coeficientes de regressão apresentados foram estatisticamente significantes, $p < .05$.

²¹ Os índices de ajuste do modelo utilizados na Análise Fatorial Confirmatória e na Modelagem de Equações Estruturais são os mesmos. Na realidade, a AFC pode ser considerada um tipo especial de MEE.

Como indicado na FIG. 6, a definição “Formal/controle do professor” foi predita positivamente pela concepção “Melhora (aluno)”, $\beta = .15$ e pela concepção “Emoção (sala)”, $\beta = .23$. Essa mesma definição foi predita negativamente pela concepção “Emoção (aluno)”, $\beta = -.56$ e pela concepção “Irrelevante (ruim)”, $\beta = -.39$. ($R^2 = .11$)

A definição “Formal/controle do aluno” foi predita positivamente pela concepção “Emoção (sala)”, $\beta = .28$ e negativamente pelas concepções “Emoção (aluno)”, $\beta = -.49$ e “Irrelevante (ruim)”, $\beta = -.22$. ($R^2 = .08$)

A definição “Informal/controle do professor” foi predita positivamente pela concepção “Emoção (sala)”, $\beta = .20$ e negativamente pela concepção “Irrelevante (ignorar)”, $\beta = -.22$. ($R^2 = .11$)

Esse modelo revelou que 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/controle do professor” ainda aumenta 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 o 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. Nos resultados encontrados na Nova Zelândia, somente práticas avaliativas informais foram preditas positivamente pela concepção “Emoção-sala” (BROWN *et al.*, 2009).

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.*, 2009). Assim, a hipótese 4 foi parcialmente corroborada.

O sentido da relação entre a definição “informal” e a concepção “Irrelevante-ignorar” na amostra brasileira é o resultado mais difícil de ser interpretado do ponto de vista teórico. Para que isso aconteça, seria necessário dizer que avaliações informais como a autoavaliação estivessem sendo realizadas no ensino superior e levadas em consideração pelo professor. Para entender melhor esse resultado no Brasil, seria necessário descobrir em pesquisas futuras como atividades avaliativas informais estão sendo conduzidas no ensino superior. Entretanto, apesar dos alunos da amostra brasileira não ignorarem esse tipo de avaliação, um ponto importante que discutimos na análise do Escalonamento Multidimensional precisa ser lembrado: quando os alunos definem o que realmente é avaliação, eles na verdade não consideram práticas de avaliação informais (ex: avaliação feita pelos colegas). Os alunos associaram a avaliação essencialmente com práticas formais.

Outro ponto relevante da análise do Escalonamento Multidimensional a ser lembrado aqui é que nove dos 12 tipos de atividades avaliativas caíram em quadrantes “controlados pelo professor”. Após a AFC, esse número aumentou para 10, pois o item “D” migrou de “Informal/controle do aluno” para “Informal/controle do professor”. Dessa forma, nossa hipótese é que os alunos da amostra brasileira não ignoram atividades informais por perceberem que elas são controladas pelo professor. Isso aumenta a probabilidade dessas atividades serem utilizadas de forma relevante. Essa interpretação vai na mesma direção de estudos anteriores que demonstram que os alunos expressam uma necessidade ou desejo de um forte direcionamento por parte dos professores (PETERSON; IRVING, 2008). Isso também fica evidente nos nossos resultados de outra forma: se a atividade é formal e controlada pelo professor, os alunos acreditam que ela ajuda a melhorar o aprendizado deles. Porém, o nosso modelo demonstra que, no pensamento dos estudantes, atividade formal e controlada pelo aluno não melhora a aprendizagem deles.

Outra possível explicação para os estudantes não ignorarem atividades informais se refere a um aspecto de nossa amostra, que era em sua maioria constituída por alunos de um Centro Universitário privado (Pública, $N= 297$, Privada, $N= 459$). Essa IES privada havia passado recentemente por uma reforma curricular onde foi introduzido em

todos os cursos um trabalho interdisciplinar. Esse tipo de trabalho faz uso de mais atividades avaliativas informais e mais uso de avaliação formativa. Essa é uma hipótese a ser considerada, pois acreditamos que nossas concepções sobre algo surgem de experiências concretas no nosso cotidiano. Portanto, é possível que algumas práticas estejam começando a se modificar no ensino superior.

6.6 - Considerações finais

Analizando os resultados da pesquisa de uma maneira global, temos um duplo desafio: mesmo que os professores estejam preparados para trabalhar com uma variedade de práticas avaliativas (o que não corresponde à realidade), nossos dados e nossa prática como docente nos levam a pensar que seria necessário também um trabalho de “convencimento” dos alunos de que tais práticas podem contribuir com seu aprendizado. Um exemplo disso foi a pesquisa de Vieira (2006), citada na revisão da literatura, que investigou as representações de estudantes sobre o portfólio como instrumento avaliativo. Os resultados indicaram que os alunos ainda precisavam ser persuadidos de que o portfólio é uma estratégia de avaliação legítima. Esse mesmo problema se apresenta com relação a outros tipos de avaliação interativo-informais, como a avaliação feita pelos colegas. Na nossa prática como docente, os alunos geralmente apresentam resistência em ser avaliados pelos seus pares, resistência em fazer atividades não pontuadas, etc. (merece destaque aqui a questão de que a concepção “Melhora” na nossa pesquisa só é associada no pensamento dos alunos com a definição de avaliação como uma prática formal e controlada pelo professor). A justificativa dos estudantes normalmente se baseia no fato deles não confiarem muito em outros alunos porque eles não saberiam o suficiente para comentar sobre o que é necessário ser melhorado. Nesse sentido, o próprio professor poderia contribuir para diminuir essa resistência. Uma alternativa possível para todas as ocasiões em que os próprios alunos forem avaliar os colegas seria o docente explicitar os critérios, mostrar exemplos de correções para os alunos. Outra opção é comparar a avaliação do professor e dos alunos e discutir possíveis diferenças. Essas opções são viáveis para várias atividades avaliativas. Portanto, inovar no planejamento de unidades didáticas e inovar na avaliação são atividades inseparáveis, inter-relacionadas. Para mudar a prática educativa é necessário mudar a prática da avaliação. Isso implicaria em uma mudança nos valores dos professores (VESLIN, 1992).

Pode-se dizer que a inclusão do SCoA em pesquisas que investigam o desempenho cognitivo dos alunos é uma alternativa possível. Nas ocasiões em que o

questionário não puder ser aplicado na sua totalidade, recomenda-se a utilização de pelo menos os itens da concepção “Melhora”, que parece estar alinhada com uma resposta positiva de auto-regulação do aprendizado (BROWN; PETERSON; IRVING, 2009). 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; as definições de avaliação incluindo novos tipos de atividades avaliativas; a relação entre as concepções e as definições de avaliação. Além disso, outros estudos devem ser realizados com o intuito de comparar também os resultados empíricos que encontramos.

Por fim, nosso trabalho contribuiu para suprir uma lacuna existente nas pesquisas educacionais brasileiras, inclusive nas pesquisas sobre as concepções de avaliação dos alunos: 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.

REFERÊNCIAS

- AFONSO, A. J. *Avaliação educacional regulação e emancipação: para uma sociologia das políticas avaliativas contemporâneas*. 2. ed. São Paulo: Cortez, 2002. 151 p.
- ALDRIDGE, J. M.; FRASER, B. J.; HUANG, T. C. I. Investigating classroom environments in Taiwan and Australia with multiple research methods. *Journal of Educational Research*, v. 93, p. 48-62, 1999.
- ATKINSON, P. *Assessment 5-14: What do pupils and parents think?* (Spotlight No. 87). Edinburgh, UK: The SCRE Centre, University of Glasgow, 2003.
- BABBIE, E. R. *Métodos de pesquisas de survey*. Belo Horizonte: Ed. UFMG, 1999. 519p.
- BANDURA, A. Social cognitive theory: an agentic perspective. *Annual review of psychology*, v. 52, p. 1-26, 2001.
- BANDURA, A. *Social Foundations of thought and Action - A social cognitive theory*. Englewood Cliffs: Prentice Hall, 1986.
- BARDIN, L. *Análise de conteúdo*. Lisboa: Edições 70, 1977.
- BARRETT, P. Structural equation modeling: Adjudging model fit. *Personality and Individual Differences*, v. 42, n. 5, p. 815-824, 2007.
- BLAIKIE, F.; SCHONAU, D.; STEERS, J. Preparing for portfolio assessment in art and design: a study of opinions and experiences of exiting secondary school students in Canada, England, and The Netherlands. *The International Journal of Art & Design Education*, v. 23, n. 3, p. 302-315, 2004.
- BOCLIN, R. Avaliação de docentes do ensino superior: um estudo de caso. *Ensaio: aval. pol. públ. Educ.*, Rio de Janeiro, v. 12, n. 45, 2004. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S010440362004000400004&lng=pt&nrm=iso>. Acesso em: 02 Nov 2007.
- BONAMINO, A. C. *Tempos de avaliação educacional: o SAEB, seus agentes, referências e tendências*. Rio de Janeiro: Quartet, 2002. 192p.
- BORGES, M. N.; GONÇALVES, M. C. N. S.; CUNHA, F. M. Teaching and learning conceptions in Engineering Education: an innovative approach on Mathematics. *European Journal of Engineering Education*, v. 28, n. 4, p. 523-534, Dez. 2009.
- BORUCHOVITCH, E. A auto-regulação da aprendizagem e a escolarização inicial. In: BORUCHOVITCH, E.; BZUNECK, J. A. (orgs.) *Aprendizagem: processos psicológicos e o contexto social na escola*. Petrópolis: Vozes, 2004. cap. 2, p. 55-88.
- BRASIL. Congresso Nacional. *Lei n. 10.861*, de 14 de abril de 2004. Institui o Sistema Nacional de Avaliação da Educação Superior - SINAES e dá outras Providências. Brasília, DF. Disponível em: <http://www.inep.gov.br/superior/avaliacao_institucional/legislacao.htm>. Acesso em: 07 dez. 2009.

BRASIL. Congresso Nacional. *Decreto Federal n.º 5.773*, de 9 de maio de 2006. Dispõe sobre o exercício das funções de regulação, supervisão e avaliação de instituições de educação superior e cursos superiores de graduação e sequenciais no sistema federal de ensino. Brasília, DF. Disponível em: < http://www.inep.gov.br/download/superior/institucional/Decreto_5773_06.pdf>. Acesso em: 07 dez. 2009.

BRASIL. Congresso Nacional. *Decreto n. 3.860*, de 9 de julho de 2001. Dispõe sobre a organização do ensino superior, a avaliação de cursos e instituições, e dá outras providências. Brasília, DF. Disponível em: <http://www.inep.gov.br/superior/avaliacao_institucional/legislacao.htm>. Acesso em: 07 dez. 2009.

BRASIL. Congresso Nacional. *Lei n. 9.131*, de 24 de novembro de 1995. Altera dispositivos da Lei nº 4.024, de 20 de dezembro de 1961, e dá outras providências. Brasília, DF. Disponível em: < <http://www.planalto.gov.br/ccivil/leis/L9131.htm>>. Acesso em: 09 dez. 2009.

BRASIL. Congresso Nacional. *Lei n. 9.394*, de 20 de dezembro 1996. Lei de Diretrizes e Bases da Educação Nacional. Estabelece as diretrizes e bases da educação nacional. Disponível em: <http://www.planalto.gov.br/ccivil_03/Leis/L9394.htm>. Acesso em: 10 dez. 2009.

BRASIL. Constituição (1988). *Constituição da República Federativa do Brasil*, 1988. Brasília, DF, 5 out. 1988. Disponível em: < http://www.planalto.gov.br/ccivil_03/constituicao/constitui%C3%A7ao.htm>. Acesso em: 10 dez. 2009.

BRASIL. Ministério da Educação. *Portaria n. 2.051*, de 9 de julho de 2004. Regulamenta os procedimentos de avaliação do Sistema Nacional de Avaliação da Educação Superior (SINAES), instituído na Lei no 10.861, de 14 de abril de 2004. Brasília, DF. Disponível em: < http://www.inep.gov.br/superior/avaliacao_institucional/legislacao.htm>. Acesso em: 10 dez. 2009.

BRASIL. Ministério da Educação. *Portaria normativa n. 12*, de 5 de setembro de 2008. Institui o Índice Geral de Cursos da Instituição de Educação Superior (IGC). Brasília, DF. Disponível em: <http://www.inep.gov.br/download/condicoes_ensino/2008/PORTARIA_NORMATIVA_12.pdf>. Acesso em: 10 dez. 2009.

BROWN, G. T. L. *Conceptions of Assessment: Understanding What Assessment Means to Teachers and Students*. New York: Nova Science Publishers, 2008. 198 p.

BROWN, G. T. L. Measuring attitude with positively packed self-report ratings: Comparison of agreement and frequency scales. *Psychological Reports*, v. 94, p. 1015-1024, 2004.

BROWN, G. T. L. *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.

BROWN, G. T. L. *Students' conceptions of assessment (SCoA) inventory* (Versions 1-6). Unpublished test. Auckland, NZ: University of Auckland. 2003.

BROWN, G. T. L. *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.

BROWN, G. T. L.; HIRSCHFELD, G. H. F. Students' conceptions of assessment and mathematics achievement: Evidence for the power of self-regulation. *Australian Journal of Educational and Developmental Psychology*, 7, p. 63-74, 2007.

BROWN, G. T. L.; HIRSCHFELD, G. H. F. 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. Beliefs that make a difference: Students' conceptions of assessment and academic performance. In: 6TH BIENNIAL CONFERENCE OF THE INTERNATIONAL TEST COMMISSION, 2008, Liverpool, UK. *Anais da 6th Biennial Conference of the International Test Commission*. Liverpool: 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, 2009.

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. (Eds.) *Student perspectives on assessment: What students can tell us about improving school outcomes*. Greenwich, CT: Information Age Press, 2009.

BRYANT, F. B.; YARNOLD, P. R. Principal-Components Analysis and Exploratory and Confirmatory Factor Analysis. In: GRIMM, L. G.; YARNOLD, P. R. (Eds.), *Reading and Understanding Multivariate Statistics*. Washington, DC: APA, 2000. cap. 4, p. 99-136.

BZUNECK, J. A. As crenças de auto-eficácia e o seu papel na motivação do aluno. In: BORUCHOVITCH, E.; BZUNECK, J. A. (orgs.) *A motivação do aluno: contribuições da psicologia contemporânea*. Petrópolis: Vozes, 2001. cap. 6, p.116-133.

CACIONE, C.; SOUZA, N. A. Avaliação da aprendizagem: desvelando concepções de licenciandos em música. In: XV CONGRESSO DA ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM MÚSICA (ANPPOM), 2005. *Anais do XV Congresso da Associação Nacional de Pesquisa e Pós-Graduação em Música (ANPPOM)*, 2005. p. 650-658.

CALDERÓN, A. I. Universidades Mercantis: a institucionalização do mercado universitário em questão. *São Paulo em Perspectiva*, v. 14, n.1, p. 61-72, 2000.

CAMARGO, A. L. C. O discurso sobre a avaliação escolar do ponto de vista do aluno. *Rev. Fac. Educ.*, v. 23, n. 1-2, Jan./Dez. 1997.

CHAUI, M. A universidade pública sob nova perspectiva. *Revista Brasileira de Educação*, n. 24, p. 5-15, Set /Out /Nov /Dez. 2003.

CHEUNG, G. W.; RENSVOOLD, R. B. Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, v. 9, n. 2, p. 233-255, 2002.

CLARK, C.; PETERSON, P. Teachers' thought processes. In: WITTROCK, M. (Ed.), *Handbook of research on teaching*. New York: MacMillan, 1986. p. 255-296.

CLAUDIO, V. Domínios de esquemas precoces na depressão. *Aná. Psicológica*, v.27, n.2, p.143-157, Jun. 2009. Disponível em: <http://www.scielo.oces.mctes.pt/scielo.php?script=sci_arttext&pid=S0870-82312009000200003&lng=pt&nrm=iso>. Acesso em: 15 jul. 2010.

COMREY, A. L.; MONTAG, I. Comparison of factor analytic results with two choice and seven choice personality item formats. *Applied Psychological Measurement*, v. 6, n. 3, p. 285-289, 1982.

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, Inc., 2000. 464 p.

DORIA, C.; TUBINO, M. J. G. Avaliação da busca da cidadania pelo Projeto Olímpico da Mangueira. *Ensaio: aval. pol. públ. Educ.*, Rio de Janeiro, v. 14, n. 50, 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-40362006000100006&lng=pt&nrm=iso>. Acesso em: 02 Nov 2007.

DUFFIELD, K. E.; SPENCER, J. A. A survey of medical students' views about the purposes and fairness of assessment. *Medical Education*, v. 36, p. 879-886, 2002.

DUROZOI, G.; ROUSSEL, A. *Dicionário de filosofia*. Campinas, São Paulo: Papirus, 1996. 511p.

ENTWISTLE, N. J.; ENTWISTLE, A. Contrasting forms of understanding for degree examinations: The student experience and its implications. *Higher Education*, 22, p. 205-227, 1991.

ETHELL, R. G.; MCMENIMAN, M. M. A critical first step in learning to teach: Confronting the power and tenacity of student teachers' beliefs and preconceptions. In: SURGRUE, C.; DAY, C. (Eds.). *Developing teachers and teacher practice: International research perspectives*. London: Routledge/Falmer Press, 2002. p. 216-233.

FALCHIKOV, N.; BOUD, D. Student self-assessment in higher education: a meta-analysis. *Review of Educational Research*, v. 59, n. 4, p. 395-430, 1989.

FALCHIKOV, N.; GOLDFINCH, J. Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Review of Educational Research*, v.70, n. 3, p. 287-322, 2000.

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.

FELIX, G. T. *Reconfiguração dos modelos de universidade pelos formatos de avaliação: efeitos no Brasil e Portugal*. 2008. 295 f. Tese (Doutorado em Educação) - Faculdade de Educação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.

FIELD, A. *Discovering statistics using SPSS*. Thousand Oaks, CA: Sage Publications, 2005. 779 p.

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.

FOWLER, F. J. *Survey Research Methods*. Newbury Park, CA: Sage, 1993.

FRASER, B. J. Classroom environment instruments: Development, validity and applications. *Learning Environments Research*, v. 1, p. 7-33, 1998.

GABLE, R. K.; WOLF, M. B. *Instrument development in the affective domain: measuring attitudes and values in corporate and school settings*. Boston: Kluwer Academic Publishers, 1993.

GALL, M. D.; GALL, J. P.; BORG, W. R. *Educational research: an introduction*. New York: Pearson, 2007. 672 p.

GATTI, B. A. Estudos quantitativos em educação. *Educ. Pesqui.*, São Paulo, v. 30, n. 1, Abr. 2004 . Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S151797022004000100002&lng=pt&nrm=iso>. Acesso em: 18 jul. 2010.

GESSER, V. *Avaliação institucional da universidade: qual seu significado para os membros de uma instituição*. 1996. Dissertação (Mestrado em Educação) - Faculdade de Educação, Pontifícia Universidade Católica de São Paulo, São Paulo, 1996.

GUIMARÃES, M. H. O. *Avaliação: impactos sobre o ensino superior*. Belo Horizonte: FUMEC-FACE, C/ Arte, 2003. 224 p.

HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. *Análise multivariada de dados*. 5. ed. Porto Alegre: Bookman, 2005. 593 p.

HAMBLETON, R. K.; PATSULA, L. Increasing the validity of adapted tests: myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, v. 1, n. 1, p. 1-30, 1999.

HARKNESS, J. A.; SCHOUA-GLUSBERG, A. Questionnaires in Translation. In: HARKNESS, J. A.; VAN DE VIJVER, F. J. R.; MOLHER, P. P. (Eds.) *Cross-Cultural Survey Methods*. New Jersey: John Wiley & Sons, 1998. p. 87- 127.

HARRIS, L. R.; BROWN, G. T. L.; HARNETT, J. A. Assessment from students' perspectives: Using pupil drawings to examine their conceptions of assessment. In: MCINERNEY, D. M.; BROWN, G. T. L.; LIEM, G. A. D. (Eds.) *Student perspectives*

on assessment: What students can tell us about improving school outcomes. Greenwich, CT: Information Age Press, 2009.

HAWES, E. M. Assessment in a pre-service teacher education programme: The rhetoric and the practice of standards-based assessment. *Asia-Pacific Journal of Teacher Education*, v. 30, n. 1, p. 93-106, 2002.

HIRSCHFELD, G. H. F.; BROWN, G. T. L. Students' conceptions of assessment: Factorial and structural invariance of the SCoA across sex, age, and ethnicity. *European Journal of Psychological Assessment*, v. 25, n. 1, p. 30-38, 2009.

HOYLE, R. H. *Structural equation modeling: Concepts, issues, and applications.* Thousand Oaks, CA: Sage Publications, 1995.

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.

HUGHES, J. *The philosophy of social research.* London: Longman Group UK, 1980.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Ações Internacionais - PISA. Disponível em: <<http://www.inep.gov.br/internacional/novo/PISA/oquee.htm>>. Acesso em: 07 jul. 2010.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Avaliação dos Cursos de Graduação. Disponível em: <<http://www.inep.gov.br/superior/condicoesdeensino/>>. Acesso em: 08 jul. 2010.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Censo da Educação Superior. Disponível em: <<http://www.inep.gov.br/superior/censosuperior/sinopse/default.asp>>. Acesso em: 16 jul. 2010.

JORBA, J.; SANMARTÍ, N. A função pedagógica da avaliação. In: BALLESTER, M. (Org.) *Avaliação como apoio à aprendizagem.* Porto Alegre: Artmed, 2003. cap 2, p.23-45.

KLEM, L. Structural equation modeling. In: GRIMM, L. G.; YARNOLD, P. R. (Eds.), *Reading and Understanding More Multivariate Statistics.* Washington, DC: APA, 2000. cap. 7, p. 227-260.

LAM, T. C. M.; KLOCKARS, A. J. Anchor point effects on the equivalence of questionnaire items. *Journal of Educational Measurement*, v. 19, n. 4, p. 317-322, 1982.

LEITE, D.; SANTIAGO, R.; SARRICO, C.; LEITE, C.; POLIDORI, M. Students' perceptions on the influence of institutional evaluation on universities. *Assessment & Evaluation in Higher Education*, v. 31, n. 6, p. 625-638, Dec. 2006.

LEMOES, L. S. *A identidade do licenciado em Psicologia: uma análise das regulamentações brasileiras entre 1962 e 2007.* 2008. 113 f. Dissertação (Mestrado em

Educação) - Faculdade de Educação, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

LOPES, M. G. *Concepções de licenciandos sobre ensino e aprendizagem: análise de uma disciplina de Prática de Ensino de Ciências Biológicas*. 2007. 155 f. Dissertação (Mestrado em Educação) - Faculdade de Educação, Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

MACHADO, L. M.; MAIA, G. Z. A.; LABEGALINI, A. C. F. B. *Pesquisa em Educação: passo a passo*. Marília: Edições M3T Tecnologia e Educação, 2007. 176 p.

MACHADO, S. M. G. *Concepções e práticas - o dilema da avaliação da aprendizagem: um estudo de caso da prática avaliativa de professores da rede estadual de ensino do Maranhão*. 2006. 217 f. Dissertação (Mestrado em Educação) - Faculdade de Educação, Universidade Federal do Maranhão, São Luís, 2006.

MARSH, H. W.; HAU, K.; GRAYSON, D. Goodness of fit in structural equation models. In: MAYDEU-OLIVARES, A.; MCARDLE, J. J. (Eds.), *Contemporary Psychometrics: A Festschrift for Roderick P. McDonald*. Mahwah, NJ: Lawrence Erlbaum Associates, 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 percepção dos alunos do comportamento comunicativo do professor de ciências*. 2006. 170 f. Dissertação (Mestrado em Educação) - Faculdade de Educação, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

MATOS, D. A. S.; CIRINO, S. D.; LEITE, W. L. Instrumentos de avaliação do ambiente de aprendizagem da sala de aula: uma revisão da literatura. *Ensaio. Pesquisa em Educação em Ciências*, Belo Horizonte, v. 10, n.1, p. 117-130, Jul. 2008.

MEZZAROBBA, L. Concepções de Avaliação de Professores e Alunos de Farmácia e Bioquímica da Universidade Estadual de Londrina, Paraná. *Revista Brasileira de Educação Médica*, Rio de Janeiro, v. 24, n. 3, p. 53-61, Out./Dez. 2000.

MINAYO, M. C. S.; SANCHES, O. Quantitativo-qualitativo: oposição ou complementariedade? *Caderno de Saúde Pública*, Rio de Janeiro, v. 9, n. 3, p. 239-262, Jul – set. 1993.

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.

MUTHÉN, L. K.; MUTHÉN, B. O. *Mplus* (Version 4.2). Los Angeles, CA: Muthén & Muthén, 2006.

NASCIMENTO, E.; FIGUEIREDO, V. L. M. WISC-III e WAIS-III: Alterações nas Versões Originais Americanas Decorrentes das Adaptações para Uso no Brasil. *Psicologia: Reflexão e Crítica*, v. 15, n. 3, p. 603-612, 2002.

NEVES, M. S. *Processo discursivo e subjetividade: vozes preponderantes na avaliação da oralidade em língua estrangeira no ensino universitário*. 2002. 259 f. Tese (Doutorado em Linguística Aplicada) - Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, São Paulo, 2002.

OFIR, C.; REDDY, S. K.; BECHTEL, G. G. Are semantic response scales equivalent? *Multivariate Behavioral Research*, v. 22, n. 1, p. 21-38, 1987.

ORY, J. C.; WISE, S. L. Attitude change measured by scales with 4 and 5 response options. In: ANNUAL MEETING OF THE NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION, 1981, Chicago. *Anais do Annual Meeting of the National Council on Measurement in Education*. Chicago: 1981.

PAJARES, F.; OLAZ, F. Teoria social cognitiva e auto-eficácia: uma visão geral. In: BANDURA, A.; AZZI, R. G.; POLYDORO, S. A. J. *Teoria social cognitiva: conceitos básicos*. Porto Alegre: Artmed, 2008. cap. 4, p. 97-114.

PAJARES, M. F.; GRAHAM, L. Formalist thinking and language arts instruction: Teachers' and students' beliefs about truth and caring in the teaching conversation. *Teaching & Teacher Education*, v. 14, n. 8, p. 855-870, 1998.

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

PASQUALI, L. *Psicometria*. Teoria dos testes na Psicologia e na Educação. Petrópolis: Vozes, 2003.

PELLISSON, J. A. *Percepções de duas professoras de língua estrangeira (inglês) e de seus alunos sobre avaliação: implicações para a formação do professor*. 2007. 154 f. Dissertação (Mestrado em Linguística Aplicada) - Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2007.

PERDIGÃO, A. L. R. V. Concepções pessoais de futuros professores sobre processos de aprendizagem e de ensino. In: REALI, A.M.M.R.; MIZUKAMI, M.G.N. (orgs.) *Formação de professores: práticas pedagógicas e escola*. São Carlos: EdUFSCar, 2002. cap. 11, p. 265-291.

PEREIRA, M. S. F. *Formação de professores e avaliação: um estudo da percepção dos alunos de um curso de pedagogia*. 2006. 159 f. Dissertação (Mestrado em Educação) - Faculdade de Educação, Universidade Estadual de Campinas, Campinas, 2006.

PETERSON E. R.; IRVING S. E. Secondary school students' conceptions of assessment and feedback. *Learning and Instruction*, v. 18, n. 3, p. 238-250, 2008.

PIAGET, J. *Seis estudos de psicologia*. Rio de Janeiro: Forense-Universitária, 1990.

POLYDORO, S. A. J.; AZZI, R. G. Auto-regulação: aspectos introdutórios. In: BANDURA, A.; AZZI, R. G.; POLYDORO, S. A. J. *Teoria social cognitiva: conceitos básicos*. Porto Alegre: Artmed, 2008. cap. 7, p. 149-164.

RAVAGNANI, F. R.; CATELAN, L. *Glossário de estatística*. São Paulo: Netra, 2002.

REAY, D.; WILIAM, 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.

REIS, C. B. *O uso dos resultados do Exame Nacional de Desempenho dos Estudantes nos cursos de física da Universidade de Brasília e da Universidade Católica de Brasília (2005)*. 2009. 171 f. Dissertação (Mestrado em Educação) - Faculdade de Educação, Universidade de Brasília, Brasília, 2009.

RUSS, J. *Dicionário de filosofia*. São Paulo: Scipione, 1994. 382p.

SAMBELL, K.; MCDOWELL, L. The construction of the hidden curriculum: Messages and meanings in the assessment of student learning. *Assessment and Evaluation in Higher Education*, v. 23, n. 4, p. 391-402, 1998.

SAMPAIO, H. *O setor privado de ensino superior no Brasil*. 1998. Tese (Doutorado em Ciência Política) - Faculdade de Filosofia Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 1998.

SANT'ANNA, I. M. *Por que avaliar? Como avaliar? Critérios e instrumentos*. Petrópolis: Vozes, 1995.

SAVIANI, D. Sobre a natureza e especificidade da educação. *Em Aberto*, Brasília, n. 22, Jul/ago. 1984.

SCHACTER, D. L. The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, v. 54, n. 3, p. 182-203, 1999.

SCHEDLER, A. Conceptualizing Accountability. In: SCHEDLER, A.; DIAMOND, L.; PLATTNER, M. F. *The Self-Restraining State: Power and Accountability in New Democracies*. London: Lynne Rienner Publishers, 1999. p. 13-28.

SGUISSARDI, V. Universidade pública estatal: entre o público e o privado/mercantil. *Educ. Soc.*, Campinas, v. 26, n. 90, p. 191-222, Jan./Abr. 2005.

SILVA, G. J. C. *O ensino superior privado: o conflito entre lucro, expansão e qualidade*. 2008. 259 f. Tese (Doutorado em Sociologia) - Instituto de Ciências Sociais, Universidade de Brasília, Brasília, 2008.

SKINNER, B. F. *Sobre o Behaviorismo*. 9. ed. São Paulo: Cultrix, 1993.

SOARES, J. F.; ALVES, M. T. G.; OLIVEIRA, R. M. O efeito de 248 escolas de nível médio no vestibular da UFMG nos anos de 1998, 1999 e 2000. *Estudos em Avaliação Educacional*, n. 24, p. 69-117, Jul-Dez. 2001.

SOBRINHO, J. D. Avaliação da educação superior: elementos para análise e proposta. In: MALAVAZI, M. M. S. (Org.) *Avaliação: desafio dos novos tempos*. São Paulo: Komedi, 2006. cap. 1, p. 15-42.

SOUZA, D. B.; VASCONCELOS, M. C. C. Os Conselhos Municipais de Educação no Brasil: um balanço das referências nacionais (1996-2002). *Ensaio: aval.pol.públ.Educ.*, Rio de Janeiro, v. 14, n. 50, mar. 2006. Disponível em:

<http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-40362006000100004&lng=pt&nrm=iso>. Acesso em: 02 Nov 2007.

STALANS, L. J. Multidimensional scaling. In: GRIMM, L. G.; YARNOLD, P. R. (Eds.), *Reading and Understanding Multivariate Statistics*. Washington, DC: APA. 1995. cap. 5, p. 137-168.

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. (Eds.), *Reading and Understanding More Multivariate Statistics*. Washington, DC: APA, 2000. cap. 8, p. 261-283.

VASCONCELLOS, M. M. M.; OLIVEIRA, C. C.; BERBEL, N. A. N. O professor e a boa prática avaliativa no ensino superior na perspectiva de estudantes. *Interface - Comunic, Saúde, Educ*, Botucatu, v. 10, n. 20, p.443-456, Jul/dez. 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1414-32832006000200012&lng=pt&nrm=iso>. Acesso em: 07 Janeiro 2008.

VEIGA, I. P. A.; RESENDE, L. M. G.; FONSECA, M. Aula universitária e inovação. In: VEIGA, I. P. A.; CASTANHO, M. E. (Orgs.) *Pedagogia universitária: a aula em foco*. Campinas: Papirus, 2000. p.161-191.

VERHINE, R. E.; DANTAS, L. M. V.; SOARES, J. F. Do Provão ao ENADE: uma análise comparativa dos exames nacionais utilizados no Ensino Superior Brasileiro. *Ensaio: aval. pol. públ. Educ.*, Rio de Janeiro, v. 14, n. 52, 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S010440362006000300002&lng=pt&nrm=iso>. Acesso em: 02 Nov 2007.

VESLIN, O. J. *Corriger des copies*. Paris: Hachette Éducation, 1992.

VIEIRA, R. L. B. *Política pública de avaliação da Educação Superior - o SINAES na Universidade Pública Estadual: o caso da Universidade do Estado da Bahia – UNEB*. 2008. 163 f. Dissertação (Mestrado em Políticas Sociais e Cidadania) – Universidade Católica do Salvador, Salvador, 2008.

VIEIRA, V. M. O. *Representações sociais e avaliação educacional: o que revela o portfolio*. 2006. 261 f. Tese (Doutorado em Psicologia da Educação) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2006.

VYGOTSKY, L. S. *A formação social da mente*. São Paulo: Martins Fontes, 1990.

VYGOTSKY, L. S. *Pensamento e linguagem*. São Paulo: Martins Fontes, 1993.

WALKER, S. L. *Learning environment research: A review of the literature* (Learning Environments Monograph No. 2). San Marcos, TX: Texas State University – San Marcos, Geography Department, 2004. Disponível em: <<http://uweb.txstate.edu/~sw36/monographs>>. Acesso em: 14 ago. 2005.

WEISZ, T. As idéias, concepções e teorias que sustentam a prática de qualquer professor, mesmo quando ele não tem consciência delas. In: WEISZ, T. *O diálogo entre o ensino e a aprendizagem*. São Paulo: Ática, 1999. p. 55-63.

WERLE, F. O. C. Gestão da educação municipal: composição dos Conselhos Municipais de Educação do Rio Grande do Sul. *Ensaio: aval. pol. públ. Educ.*, Rio de Janeiro, v. 14, n. 52, 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S010440362006000300005&lng=pt&nrm=iso>. Acesso em: 02 Nov 2007.

WHEELOCK, A.; BEBELL, D.; HANEY, W. What can student drawings tell us about high-stakes testing in Massachusetts? *Teachers College Record*, ID Number: 10634, 2000.

WHITE, R. T. Conceptual and conceptional change. *Learning and Instruction*, v. 4, p. 117-121, 1994.

ZEIDNER, M. Key facets of classroom grading: A comparison of teacher and student perspectives. *Contemporary Educational Psychology*, v. 17, p. 224-243, 1992.

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

ANEXO A - Questionário *Students' Conceptions of Assessment* (SCoA - VI)

Por favor preencha as informações a seguir:

Sexo: () Masculino () Feminino Idade: _____

Curso: _____ Período: _____

Todos os dados coletados são confidenciais.

Leia atentamente as instruções a seguir:

Esse questionário contém afirmações sobre experiências relativas à avaliação. Para respondê-lo, leve em consideração as suas experiências de avaliação no Ensino Superior de um modo geral. Pense sobre a melhor maneira na qual cada afirmação descreve o que a avaliação é para você. Para cada afirmação, você tem 6 opções de resposta: Discordo fortemente; Discordo na maior parte; Concordo ligeiramente; Concordo moderadamente; Concordo na maior parte; Concordo fortemente. Para cada item, escolha somente uma dessas opções. Marque com um X a resposta escolhida.

Não existem respostas “certas” ou “erradas”. Isso não é uma prova nem um teste. O que importa é a sua opinião. Nós queremos seu ponto de vista como aluno.

Algumas afirmações nesse questionário são parecidas com outras. Não se preocupe com isso. Simplesmente dê sua opinião sobre todas as afirmações.

Responda calmamente sem se preocupar com o tempo.

O questionário deve ser respondido individualmente e em silêncio.

Esse questionário tem 4 folhas. Confira se o seu tem 4 folhas e responda a todas. Na página 4 existe uma parte complementar para ser respondida.

(Questionário traduzido e adaptado do original em inglês *Students' Conceptions of Assessment version VI* (SCoA-VI) (BROWN, 2006), com a autorização de Gavin T L Brown da Universidade de Auckland.)

[illegible]

PARTE COMPLEMENTAR DO QUESTIONÁRIO

Quando você pensa na palavra AVALIAÇÃO, quais tipos de atividades avaliativas vêm a sua mente ?

Instruções de resposta: marque um X nas opções que quiser. Você pode marcar quantas considerar necessário.

- ☐ O professor atribui uma nota ou avalia um trabalho em grupo
- ☐ Eu atribuo uma nota ou avalio meu próprio desempenho (auto-avaliação)
- ☐ Meus colegas atribuem uma nota ou avaliam meu desempenho
- ☐ O professor me faz perguntas em voz alta na sala de aula
- ☐ O professor atribui uma nota ou avalia um trabalho escrito que eu entrego
- ☐ O professor aplica prova com consulta
- ☐ O professor me observa na sala de aula e julga o meu aprendizado
- ☐ O professor aplica prova em dupla
- ☐ O professor me dá uma nota em um trabalho feito na sala de aula
- ☐ O professor aplica prova com questões de múltipla escolha
- ☐ O professor aplica prova com questões abertas
- ☐ O professor atribui uma nota ao meu desempenho depois de se reunir ou discutir comigo sobre o meu trabalho
- ☐ Algo mais: _____ (Escreva com letra legível)
- ☐ Algo mais: _____ (Escreva com letra legível)

Você tem algum comentário ?

ANEXO B - Quadro completo do processo de tradução do SCoA

Titulação dos tradutores que participaram da tradução inversa (*back translation*):

Tradutor 1 - Graduação: Direito. Mestrado: Direito Internacional.

Tradutor 2 - Graduação: Letras. Mestrado: Linguística.

Tradutor 3 - Graduação: Psicologia. Mestrado: Psicologia Social.

Versão em inglês do questionário	Tradução para o português	Tradução inversa Tradutor 1	Tradução inversa Tradutor 2	Tradução inversa Tradutor 3	Versão final da tradução
Students' Conceptions of Assessment version VI	Concepções de avaliação de alunos versão VI	Conceptions for students' assessment version n. VI	Conceptions of assessment of students	Students concepts about evaluation version VI	Concepções de alunos sobre avaliação Versão VI
Conceptions of Assessment	Concepções de avaliação	Conceptions of assessment	Concepts of assessment	Concepts about evaluation	Concepções sobre avaliação
Strongly Disagree	Discordo fortemente	I strongly disagree	I strongly disagree	Strongly disagree	Discordo fortemente
Mostly Disagree	Discordo na maior parte	I mostly disagree	I disagree with most of it	Mostly disagree	Discordo na maior parte
Slightly Agree	Concordo ligeiramente	I slightly agree	I slightly agree	Lightly agree	Concordo ligeiramente
Moderately Agree	Concordo moderadamente	I moderately agree	I mildly agree	Moderately agree	Concordo moderadamente
Mostly Agree	Concordo na maior parte	I mostly agree	I agree with the most part	Mostly agree	Concordo na maior parte
Strongly Agree	Concordo fortemente	I strongly agree	I strongly agree	Strongly agree	Concordo fortemente
I pay attention to my assessment results in order to focus on what I could do better next time	Eu presto atenção nos meus resultados de avaliação para me concentrar no que eu posso melhorar da próxima vez	I take my assessment results into account to focus on what I can improve next time	I pay attention to the results of my assessment in order to focus on what I can improve next time	I pay attention to the results of my evaluations to concentrate in what I can improve next time.	Eu presto atenção nos meus resultados de avaliação para me concentrar no que eu posso melhorar da próxima vez

Versão em inglês do questionário	Tradução para o português	Tradução inversa Tradutor 1	Tradução inversa Tradutor 2	Tradução inversa Tradutor 3	Versão final da tradução
Assessment encourages my class to work together and help each other	A avaliação encoraja a minha turma a trabalhar junta e a ajudar uns aos outros	The assessment encourages my class to work together and to help one another	The assessment encourages my class to work together and to help each other.	The evaluation helps my class to work together and help one another	A avaliação encoraja a minha turma a trabalhar junta e a ajudar uns aos outros
Assessment is unfair to students	A avaliação é injusta com os alunos	The assessment is unfair to the students	The assessment is unfair to the students	The evaluation is unfair with the students	A avaliação é injusta com os alunos
Assessment results show how intelligent I am	Os resultados da avaliação mostram o quanto eu sou inteligente	The assessment results show how smart I am	The results of the assessment show how intelligent I am	The results of the evaluation show how smart I am	Os resultados da avaliação mostram o quanto eu sou inteligente
Assessment helps teachers track my progress	A avaliação ajuda os professores a acompanhar o meu progresso	The assessment help teachers follow up on my progress	The assessment helps the teachers to follow my development	The evaluation helps teachers to keep score of my progress	A avaliação ajuda os professores a acompanhar o meu progresso
Assessment is an engaging and enjoyable experience for me	A avaliação é uma experiência envolvente e agradável para mim	The assessment is a engaging and pleasant experience for me	The assessment is an involving and pleasant experience to me	The evaluation is a nice and involving experience for me	A avaliação é uma experiência envolvente e agradável para mim
I ignore assessment information	Eu ignoro informações da avaliação	I ignore the information from the assessment	I ignore information about the assessment	I ignore information of the evaluation	Eu ignoro as informações da avaliação
Assessment is a way to determine how much I have learned from teaching	A avaliação é uma forma de determinar o quanto eu aprendi do ensino	The assessment is a way to determine how much I learned from what was taught	The assessment is a way to determine how much I have learnt from the teaching	The evaluation is a form to determine how much I learn about what was taught	A avaliação é uma forma de determinar o quanto eu aprendi do ensino
Assessment is checking off my progress against achievement objectives or standards	A avaliação averigua o meu progresso em comparação com objetivos ou padrões de realização	The assessment evaluates my progress in comparison to the goals or patterns of accomplishment	The assessment checks my progress compared to the objectives or fulfillment patterns	The evaluation measure my progress comparing it with goals and patterns of realization	A avaliação averigua o meu progresso em comparação com os objetivos de aprendizagem

Versão em inglês do questionário	Tradução para o português	Tradução inversa Tradutor 1	Tradução inversa Tradutor 2	Tradução inversa Tradutor 3	Versão final da tradução
I make use of the feedback I get to improve my learning	Eu faço uso do <i>feedback</i> que recebo para melhorar meu aprendizado	I use the feedback I receive to enhance my learning	I make use of the feedback I get to improve my learning	I use the feedback I get to improve my learning	Eu faço uso do <i>feedback</i> que recebo para melhorar meu aprendizado
Assessment provides information on how well schools are doing	A avaliação fornece informação sobre como as escolas estão indo	The assessment gives information on how good schools are	The assessment supplies information on how the school is going	The evaluation provides information about how schools are going	A avaliação fornece informação sobre como as escolas estão indo
Assessment motivates me and my classmates to help each other	A avaliação me motiva e aos meus colegas a ajudarem uns aos outros	The assessment motivates me and my classmates to help one another	The assessment drives me and my classmates to help each other	The evaluation motivates me and my classmates to help one another	A avaliação me motiva e aos meus colegas a ajudarem uns aos outros
Assessment interferes with my learning	A avaliação interfere no meu aprendizado	The assessment interferes on my learning	The assessment interferes with my learning	The evaluation interferes in my learning	A avaliação interfere no meu aprendizado
I look at what I got wrong or did poorly on to guide what I should learn next	Eu observo o que eu fiz de errado ou de maneira insuficiente para guiar o que eu deveria aprender em seguida	I observe what I did wrong or insufficiently to guide my following learning	I observe what I have done wrong or in an insufficient manner to guide me on what I should learn next	I observe what I did wrong or insufficiently to guide what I should learn next	Eu observo o que eu fiz de errado ou de maneira insuficiente para guiar o que eu deveria aprender em seguida
I use assessments to take responsibility for my next learning steps	Eu uso as avaliações para assumir responsabilidade para os meus próximos passos de aprendizagem	I use the assessment to take on responsibility on the next steps of my learning	I use the assessments to take responsibility on my next learning steps	I use the evaluation to assume responsibility to my next learning steps	Eu uso as avaliações para assumir responsabilidade para as minhas próximas etapas de aprendizagem
Assessment results predict my future performance	Os resultados da avaliação predizem o meu desempenho futuro	The assessment results predict my future performance	The results of the assessment predict my future performance	The results of the evaluation predicts my future performance	Os resultados da avaliação predizem o meu desempenho futuro
Our class becomes more supportive when we are assessed	A nossa turma se dá mais apoio quando nós somos avaliados	Our class gives more support when we are	Our class gives itself more support when	My class gives internal more support	A nossa turma se dá mais apoio quando nós somos avaliados

		under analysis	we are assessed	when we are evaluated	
Versão em inglês do questionário	Tradução para o português	Tradução inversa Tradutor 1	Tradução inversa Tradutor 2	Tradução inversa Tradutor 3	Versão final da tradução
Teachers are over-assessing	Os professores avaliam em demasia	The teachers evaluate us in excess	Teachers assess too much	Teachers evaluate too much	Os professores avaliam exageradamente
I use assessments to identify what I need to study next	Eu uso as avaliações para identificar o que eu preciso estudar em seguida	I use the assessment to identify what I should learn in the near future	I use the assessment to identify what I need to study next	I use evaluations to identify what I need to study next	Eu uso as avaliações para identificar o que eu preciso estudar em seguida
Assessment is important for my future career or job	A avaliação é importante para a minha carreira futura ou emprego	The assessment is important for my future career or job	The assessment is important for my future career and job	The evaluation is important to my future career or job	A avaliação é importante para a minha carreira futura ou emprego
When we do assessments, there is a good atmosphere in our class	Quando nós fazemos avaliações, existe um bom clima na nossa turma	When we take evaluations there is a good atmosphere in our classroom	When we do our assessment, there is a good atmosphere in our classroom	When we do evaluations, there is a good mood in our class	Quando nós fazemos avaliações, existe um bom clima na nossa turma
Assessment results are not very accurate	Os resultados da avaliação não são muito exatos	The assessment results are not very accurate	The results of the assessment are not very precise	Results of evaluations are not very accurate	Os resultados da avaliação não são muito exatos
My teachers use assessment to help me improve	Os meus professores usam a avaliação para me ajudar a melhorar	My teachers use the assessment to help me improve	My teachers use the assessment to help me improve	My teachers use the evaluation to help me to improve	Os meus professores usam a avaliação para me ajudar a melhorar
Assessment measures the worth or quality of schools	A avaliação mede o valor ou qualidade das escolas	The assessment measures the value or quality of schools	The assessment measures the value or the quality of the schools	The evaluation measure the value or quality of the school	A avaliação mede a qualidade das escolas
Assessment makes our class cooperate more with each other	A avaliação faz a nossa turma cooperar mais uns com os outros	The assessment makes our class cooperate more to one another	The assessment makes our class cooperate more with each other	The evaluation makes our class cooperate one with the others	A avaliação faz a nossa turma cooperar mais uns com os outros
Assessment is value-less	A avaliação é sem valor	The assessment is useless	The assessment is worthless	The evaluation has no value	A avaliação é sem valor

Versão em inglês do questionário	Tradução para o português	Tradução inversa Tradutor 1	Tradução inversa Tradutor 2	Tradução inversa Tradutor 3	Versão final da tradução
Teachers use my assessment results to see what they need to teach me next	Os professores usam os meus resultados da avaliação para ver o que eles precisam me ensinar em seguida	The teachers use my assessment results to come across what they need to teach me next	The teachers use my assessment results to see what they need to teach me next	Teachers use evaluations results to see what they need to teach next	Os professores usam os meus resultados da avaliação para ver o que eles precisam me ensinar em seguida
When we are assessed, our class becomes more motivated to learn	Quando nós somos avaliados, a nossa turma se torna mais motivada para aprender	When we are evaluated our class becomes more motivated to learn	When we are assessed, our group becomes more motivated to learn	When we are evaluated our class became more motivated to learn	Quando nós somos avaliados, a nossa turma se torna mais motivada para aprender
I ignore or throw away my assessment results	Eu ignoro ou descarto os meus resultados de avaliação	I ignore or discharge my assessment results	I ignore or discard my assessment results	I ignore or trough away the results of the evaluation	Eu desconsidero os meus resultados de avaliação
Assessment shows whether I can analyse and think critically about a topic	A avaliação mostra se eu posso analisar e pensar criticamente sobre um assunto	The assessment shows me if I am able to analyze and think critically about a subject	The assessment shows if I can analyze and think critically about a subject	The evaluation shows if I can analyze and think critically about a theme	A avaliação mostra se eu posso analisar e pensar criticamente sobre um assunto
I find myself really enjoying learning when I am assessed	Eu realmente aprecio o aprendizado quando eu sou avaliado	I really appreciate the learning when I am being evaluated	I really appreciate learning when I am assessed	I really appreciate the learning when I am evaluated	Eu realmente aprecio o aprendizado quando eu sou avaliado
Assessment has little impact on my learning	A avaliação tem um impacto pequeno no meu aprendizado	The assessment has a small impact on my learning	The assessment has a small impact on my learning	The evaluation has short impact in my learning	A avaliação tem um impacto pequeno no meu aprendizado
Assessment tells my parents how much I've learnt	A avaliação diz para os meus pais o quanto eu aprendi	The assessment shows my parents how much I learned	The assessment shows my parents how much I have learnt	The evaluation shows my parents what I learned	<i>Item excluído da versão em português por questões de diferenças culturais.</i>
When you think of the word <u>ASSESSMENT</u> , which kinds or types of assessment activities come to your mind? (Fill in all that apply)	Quando você pensa na palavra <u>AVALIAÇÃO</u> , quais tipos de atividades avaliativas vêm a sua mente? (Marque um X em tudo o que se aplica)	When you think of the word <u>ASSESSMENT</u> , what kind of assessment activities come to your mind? (Write an X on every option that applies)	When you think of the word <u>assessment</u> , what types of assessment activities come to your mind? (Put an X in everything that	When you think about the word <u>evaluation</u> , what forms of evaluative activities come to your mind? (sign with an X all items that apply)	Quando você pensa na palavra <u>AVALIAÇÃO</u> , quais tipos de atividades avaliativas vêm a sua mente ? Marque um X nas opções que quiser. Você pode marcar

			applies		quantas considerar necessário.
Versão em inglês do questionário	Tradução para o português	Tradução inversa Tradutor 1	Tradução inversa Tradutor 2	Tradução inversa Tradutor 3	Versão final da tradução
An examination that takes one to three hours	Uma prova que leva de uma a três horas	A test which takes from one to three hours	A test that lasts for one to three hours	One test from one to three hours	<i>Item excluído da versão em português e substituído por outro. Ver abaixo em “Novos itens da versão em português”.</i>
I score or evaluate my own performance	Eu atribuo uma nota ou avalio meu próprio desempenho	I give myself a grade or assess my own performance	I assign a grade or I evaluate my own performance	I give myself a grade or evaluate my performance	Eu atribuo uma nota ou avalio meu próprio desempenho (auto-avaliação)
My class mates score or evaluate my performance	Meus colegas atribuem nota ou avaliam meu desempenho	My classmates give me a grade or assess my performance	My classmates assign a grade or evaluate my performance	My classmates give me a grade or evaluate my performance	Meus colegas atribuem uma nota ou avaliam meu desempenho
The teacher asks me questions out loud in class	O professor me faz perguntas em voz alta na sala de aula	The teacher asks me oral questions in the classroom	The teachers asks me oral questions in the classroom	The teacher ask me questions orally in class	O professor me faz perguntas em voz alta na sala de aula
The teacher grades or marks or scores the written work I hand in	O professor classifica ou corrige ou atribui nota ao trabalho escrito que eu entrego	The teacher classifies or corrects or marks a written work I deliver	The teacher classifies or corrects or assigns a grade to the writing work that I hand over	The teacher classify or correct or grade writing works that I turned on	O professor atribui uma nota ou avalia um trabalho escrito que eu entrego
The teacher grades me on a written test that he or she made up	O professor me dá uma nota em um teste escrito que ele fez	The teacher gives me a grade on a written test he wrote	The teacher gives me a grade on a written test he has made	The teacher grades a writing test produced by him or her	<i>Item excluído da versão em português e substituído por outro. Ver abaixo em “Novos itens da versão em português”.</i>
The teacher grades me on a written test that was written by someone other than the teacher	O professor me dá uma nota em um teste escrito que foi feito por alguém diferente do professor	The teacher gives me a grade on a written test someone else wrote	The teacher gives me a grade on a written test made by somebody else than	The teacher grades a writing test produced by another person	<i>Item excluído da versão em português por questões de diferenças culturais.</i>

			the teacher		
Versão em inglês do questionário	Tradução para o português	Tradução inversa Tradutor 1	Tradução inversa Tradutor 2	Tradução inversa Tradutor 3	Versão final da tradução
The teacher observes me in class and judges my learning	O professor me observa na sala de aula e julga o meu aprendizado	The teacher observes me in the classroom and assess my learning	The teacher observes me in the classroom and judges my learning	The teacher watch what I do in class and evaluate my learning	O professor me observa na sala de aula e julga o meu aprendizado
The teacher scores a portfolio of work I have done over the course of a term or school year	O professor atribui uma nota a um portfólio de trabalhos que eu fiz durante um semestre ou ano escolar	The teacher gives me a grade on a work portfolio I assembled during the school semester or year	The teacher assigns a grade to a portfolio with the work I have done during a semester or a school year	The teacher grade a group of works I did during the semester or school year	<i>Item excluído da versão em português por questões de diferenças culturais.</i>
The teacher scores me on an in-class written essay	O professor me dá uma nota em um trabalho feito na sala de aula	The teacher gives me a grade on a work done in the classroom	The teacher gives me a grade on a work done in the classroom	The teacher grades a work done in the class	O professor me dá uma nota em um trabalho feito na sala de aula
The teacher scores my performance after meeting or conferencing with me about my work	O professor atribui uma nota ao meu desempenho depois de se reunir ou discutir comigo sobre o meu trabalho	The teacher gives me a grade on my performance, after meeting or discussing my work with me	The teacher assigns a grade to my performance after gathering with me or discussing about my work	The teacher grades my performance after meeting or talking to me about my work	O professor atribui uma nota ao meu desempenho depois de se reunir ou discutir comigo sobre o meu trabalho
The teacher uses a checklist to judge my in-class performance	O professor usa um <i>checklist</i> para julgar o meu desempenho dentro da sala de aula	The teacher uses a checklist to assess my performance in the classroom	The teacher uses a checklist to judge my performance inside the classroom	The teacher uses a checklist to judge my performance in class	<i>Item excluído da versão em português por questões de diferenças culturais.</i>
Something else: <i>(print neatly)</i>	Algo mais: <i>(Escreva com letra legível)</i>	Something else: <i>(Write legibly)</i>	Something else: <i>(Write with a legible handwriting)</i>	Anything more? <i>(write clearly)</i>	Algo mais: <i>(Escreva com letra legível)</i>
Do you have any comments?	Você tem algum comentário?	Do you have any comments?	Do you have any comments?	Do you have any comments?	Você tem algum comentário?

Novos itens da versão em português (parte complementar)					Novos itens da versão em português (parte complementar)
<i>Item criado especificamente para a versão em português.</i>					O professor atribui uma nota ou avalia um trabalho em grupo
<i>Item criado especificamente para a versão em português.</i>					O professor aplica prova com consulta
<i>Item criado especificamente para a versão em português.</i>					O professor aplica prova em dupla
<i>Item criado especificamente para a versão em português.</i>					O professor aplica prova com questões de múltipla escolha
<i>Item criado especificamente para a versão em português.</i>					O professor aplica prova com questões abertas

ANEXO C - Roteiro de entrevista semi-estruturada sobre o SCoA

Instrução inicial:

“Leiam e respondam o questionário todo, individualmente e em silêncio. Durante a leitura, marquem a lápis qualquer coisa que vocês não entenderem bem. Depois que todos terminarem, vamos discutir em grupo as dificuldades que vocês tiveram. Se alguém terminar antes do outro, por favor aguarde em silêncio até que todos terminem.”

Perguntas:

- 1- Vocês podem dizer sobre o que o questionário fala?
- 2- Nas instruções do questionário, tem alguma coisa que vocês não entenderam?
- 3- Agora vou falar o número de cada item do questionário. Se vocês não entenderam bem algum deles, me avise para a gente discutir. Iremos conversar somente sobre os itens que vocês não entenderam.

ANEXO D - Declaração de autorização de pesquisa

Declaro para os devidos fins, que o pesquisador Daniel Abud Seabra Matos (tel. (31) 91564131), aluno do curso de Doutorado em Educação, Conhecimento e Inclusão Social da Faculdade de Educação (FaE) da Universidade Federal de Minas Gerais (UFMG), e seu orientador, Professor Dr. Sérgio Dias Cirino (tel. (31) 3409-6206), do Departamento de Métodos e Técnicas de Ensino da (FaE- UFMG) requisitaram a participação da instituição _____ denominada

_____ no estudo intitulado **“Concepções de avaliação de alunos do Ensino Superior”**.

O objetivo geral da pesquisa é analisar as concepções de avaliação de alunos do ensino superior. A participação dos alunos nesse estudo envolve a participação em entrevistas em grupo que tem como objetivos conhecer as concepções de avaliação de alunos do ensino superior e contribuir para a elaboração de um questionário. Envolve também responder um questionário de 33 itens sobre concepções de avaliação. Além disso, também é necessário que o pesquisador acesse as notas dos alunos nas disciplinas do ano atual.

Os riscos da influência negativa deste projeto na aprendizagem e no rendimento escolar são muito pequenos. Será necessário apenas um contato do pesquisador com os alunos tanto nas entrevistas quanto nos questionários. Além disso, o questionário pode ser respondido rapidamente. Portanto, os dados serão obtidos em apenas um encontro.

Entretanto, um canal de comunicação entre o pesquisador, os professores, a direção da escola e os alunos será mantido ao longo da pesquisa. Se a escola se sentir prejudicada em algum aspecto ela pode solicitar a interrupção da pesquisa a qualquer momento e por iniciativa de qualquer um dos sujeitos participantes da pesquisa.

Além disso, o nome da escola e dos demais participantes serão alterados no registro escrito. Somente os pesquisadores envolvidos neste estudo terão acesso a estas informações, que serão utilizadas apenas para fins científicos. A identidade de qualquer pessoa que participe do estudo não será revelada publicamente em nenhuma hipótese.

Os resultados desta pesquisa poderão ajudar futuramente no maior conhecimento sobre o papel das concepções de avaliação de alunos para o ensino superior, bem como nos processos de formação de professores.

Estou ciente de que tenho total liberdade para esclarecer qualquer dúvida em relação à pesquisa ou à participação da instituição, antes ou depois de minha autorização, pois todas as dúvidas deverão ser esclarecidas pelos membros da equipe responsável, pessoalmente durante a pesquisa. Também posso desligar a instituição deste projeto a qualquer momento sem necessidade de me justificar junto aos pesquisadores.

Eu, voluntariamente, autorizo a realização da pesquisa nessa instituição. Portanto, concordo com tudo que está escrito acima e dou meu consentimento.

_____, _____ de _____ de 2008.

Nome da instituição: _____.

Nome e assinatura do responsável: _____.

Contato com o **COEP/UFMG (Comitê de Ética em Pesquisa)**:
Av. Antônio Carlos, 6627, Unidade Administrativa II, 2o. andar, sala 2005.
Campus Pampulha - Belo Horizonte, MG – Brasil -31270-901
End. eletrônico: coep@prpq.ufmg.br

Telefax: (31) 3409-4592

ANEXO E - Termo de consentimento livre e esclarecido

(Para alunos que responderam ao questionário)

O pesquisador Daniel Abud Seabra Matos (tel. (31) 91564131), aluno do curso de Doutorado em Educação, Conhecimento e Inclusão Social da Faculdade de Educação (FaE) da Universidade Federal de Minas Gerais (UFMG), e seu orientador, Professor Dr. Sérgio Dias Cirino (tel. (31) 3409-6206), do Departamento de Métodos e Técnicas de Ensino da (FaE- UFMG) requisitaram minha participação neste estudo intitulado **“Concepções de avaliação de alunos do Ensino Superior”**.

O objetivo geral da pesquisa é analisar as concepções de avaliação de alunos do ensino superior. A participação neste estudo envolve responder um questionário de 33 itens e permitir que o pesquisador acesse as notas dos alunos nas disciplinas do ano atual.

Os riscos da influência negativa deste projeto na aprendizagem e no rendimento escolar são muito pequenos. Será necessário apenas um contato do pesquisador com os alunos. O questionário pode ser respondido rapidamente. Portanto, os dados serão obtidos em apenas um encontro.

Entretanto, um canal de comunicação entre o pesquisador, os professores, a direção da escola e os alunos será mantido ao longo da pesquisa. Se a escola se sentir prejudicada em algum aspecto ela pode solicitar a interrupção da pesquisa a qualquer momento e por iniciativa de qualquer um dos sujeitos participantes da pesquisa.

Além disso, o nome da escola e dos demais participantes serão alterados no registro escrito. Somente os pesquisadores envolvidos neste estudo terão acesso a estas informações, que serão utilizadas apenas para fins científicos. A identidade de qualquer pessoa que participe do estudo não será revelada publicamente em nenhuma hipótese.

Os resultados desta pesquisa poderão ajudar futuramente no maior conhecimento sobre o papel das concepções de avaliação de alunos para o ensino superior, bem como nos processos de formação de professores.

Estou ciente de que tenho total liberdade para esclarecer qualquer dúvida em relação à pesquisa ou à minha participação, antes ou depois do meu consentimento, pois todas as dúvidas deverão ser esclarecidas pelos membros da equipe responsável, pessoalmente durante a pesquisa. Também posso me desligar deste projeto a qualquer momento sem necessidade de me justificar junto aos pesquisadores.

Eu, voluntariamente, aceito participar desta pesquisa. Portanto, concordo com tudo que está escrito acima e dou meu consentimento.

_____, _____ de _____ de 2008.

Nome do aluno voluntário : _____.

Assinatura do aluno voluntário : _____.

Contato com o **COEP/UFMG (Comitê de Ética em Pesquisa)**:
Av. Antônio Carlos, 6627, Unidade Administrativa II, 2o. andar, sala 2005.
Campus Pampulha - Belo Horizonte, MG – Brasil -31270-901
End. eletrônico: coep@prpq.ufmg.br

Telefax: (31) 3409-4592

Termo de consentimento livre e esclarecido

(Para os alunos que participaram da entrevista)

O pesquisador Daniel Abud Seabra Matos (tel. (31) 91564131), aluno do curso de Doutorado em Educação, Conhecimento e Inclusão Social da Faculdade de Educação (FaE) da Universidade Federal de Minas Gerais (UFMG), e seu orientador, Professor Dr. Sérgio Dias Cirino (tel. (31) 3409-6206), do Departamento de Métodos e Técnicas de Ensino da (FaE- UFMG) requisitaram minha participação neste estudo intitulado **“Concepções de avaliação de alunos do Ensino Superior”**.

O objetivo geral da pesquisa é analisar as concepções de avaliação de alunos do ensino superior. A participação neste estudo envolve participar de uma entrevista, em um grupo de 4 alunos, sobre um questionário de 33 itens. O objetivo dessa entrevista é verificar se os alunos conseguem entender o significado dos itens.

Os riscos da influência negativa deste projeto na aprendizagem e no rendimento escolar são muito pequenos. Será necessário apenas um contato do pesquisador com os alunos. A entrevista pode ser feita rapidamente. Portanto, os dados serão obtidos em apenas um encontro.

Entretanto, um canal de comunicação entre o pesquisador, os professores, a direção da escola e os alunos será mantido ao longo da pesquisa. Se a escola se sentir prejudicada em algum aspecto ela pode solicitar a interrupção da pesquisa a qualquer momento e por iniciativa de qualquer um dos sujeitos participantes da pesquisa.

Todos os dados coletados são confidenciais. Portanto, a identidade do participante não será revelada publicamente em nenhuma hipótese, sendo o nome da instituição e dos demais sujeitos participantes alterados no registro escrito. Somente os pesquisadores envolvidos neste estudo terão acesso a estas informações, que serão utilizadas apenas para fins científicos.

Os resultados desta pesquisa poderão ajudar futuramente no maior conhecimento sobre o papel das concepções de avaliação de alunos para o ensino superior, bem como nos processos de formação de professores.

Estou ciente de que tenho total liberdade para esclarecer qualquer dúvida em relação à pesquisa ou à minha participação, antes ou depois do meu consentimento, pois todas as dúvidas deverão ser esclarecidas pelos membros da equipe responsável, pessoalmente durante a pesquisa. Também posso me desligar deste projeto a qualquer momento sem necessidade de me justificar junto aos pesquisadores.

Eu, voluntariamente, aceito participar desta pesquisa. Portanto, concordo com tudo que está escrito acima e dou meu consentimento.

_____, _____ de _____ de 2008.

Nome do aluno voluntário : _____.

Assinatura do aluno voluntário : _____.

Contato com o **COEP/UFMG (Comitê de Ética em Pesquisa)**:
Av. Antônio Carlos, 6627, Unidade Administrativa II, 20. andar, sala 2005.
Campus Pampulha - Belo Horizonte, MG – Brasil -31270-901
End. eletrônico: coep@prpq.ufmg.br

Telefax: (31) 3409-4592