

EXPLICANDO A QUALIDADE DE CURSOS SUPERIORES DE ADMINISTRAÇÃO A PARTIR DE MODELOS DE EQUAÇÕES ESTRUTURAIS

Arturo Cavalcanti Catunda *

Robert Evan Verhine **

RESUMO

Este artigo trata de uma investigação realizada para identificar fatores determinantes da qualidade de cursos da educação superior no Brasil, por meio da aplicação da técnica estatística de Modelo de Equações Estruturais (MEE). Foram utilizados dados secundários obtidos do Censo da Educação Superior e do ENADE, contemplando um universo de 2.939 cursos e 1.035.765 alunos, da área de gerenciamento e administração, do ano de 2006. Também foi utilizada a técnica estatística de Análise de Classe Latente, resultando na identificação de 11 construtos que possibilitaram elaborar três modelos, um para cada uma das variáveis dependentes: ENADE – Conceito; IDD – Conceito; e Resultado Geral no ENADE. Dos três modelos, o relacionado ao ENADE – Conceito foi o que melhor explicou a variável dependente. O uso das técnicas nesta investigação permitiu identificar o perfil de instituição de ensino superior mais relacionado ao conceito de qualidade.

Palavras-chave: Ensino superior. Eficácia no ensino. Modelo de equações estruturais.

ABSTRACT

EXPLAINING THE QUALITY OF HIGHER EDUCATION PROGRAMS OF ADMINISTRATIONS USING STRUCTURAL EQUATIONS MODEL

This article describes an investigation carried out to identify factors determining the quality of higher education programs in Brazil, applying the statistical technique of the Structural Equations Model (SEM). Secondary data were obtained from the Census of Higher Education and ENADE, covering a population of 1,035,765 students and 2939 programs, in the fields of business management and administration in 2006. We also used the statistical technique of Latent Class Analysis, resulting in identification of 11 possible constructs, that produced three models, one for each of the dependent variables: Concept – ENADE; Concept – IDD and General ENADE Results. Of the three models, the one related to Concept – ENADE best explained the dependent variable. The use of the techniques in this research allowed identifying the profile of higher education institutions more closely related to the concept of quality.

Keywords: Higher education. School effectiveness. Structural equations model.

* Doutor em Educação pela Universidade Federal da Bahia. Administrador na Petrobrás. Endereço para correspondência: Rua dos Vereadores, Condomínio Vivendas do Atlântico, Casa 69, Vila Praiana, Lauro de Freitas, Bahia. CEP: 42700-000. acatunda@yahoo.com.br

** Doutor em Educação pela Universität Hamburg. Professor Associado 3 da Faculdade de Educação da Universidade Federal da Bahia- UFBA. Endereço para correspondência: Universidade Federal da Bahia, Faculdade de Educação, Departamento de Educação I. Av. Reitor Miguel Calmon s/n, Vale do Canela, Salvador, Bahia. CEP: 40110-100. verhine@ufba.br

Introdução

Muito antes da Teoria do Capital Humano, desenvolvida por Theodore W. Schultz em meados dos anos 1950, comprovar e difundir os retornos econômicos individuais e sociais proporcionados por uma maior escolarização da população, economistas como Adam Smith (1723-1790) e Thomas Malthus (1766-1834), assim como filósofos como Rousseau (1772-1778) e Platão (428-348 a.C.) já defendiam a necessidade de o povo ser amplamente mais bem educado e preparado para conviver com as transformações ocorridas nas sociedades de suas épocas.

Entre inúmeras virtudes, a Educação é a instituição considerada capaz de moldar o ser humano; criar, fortalecer e transformar culturas; proporcionar riquezas às nações; possibilitar a transmissão dos avanços científicos por meio da instrumentalização de produtos e serviços disponibilizados à população; e dotar as pessoas de consciência crítica necessária à cidadania. Atualmente, com o acirramento da competição internacional, as transformações políticas e econômicas, a ampliação da dependência do ser humano à utilização de tecnologia cada vez mais sofisticada, as incertezas ambientais e os riscos delas decorrentes para a manutenção e melhoria do padrão de vida, a educação torna-se ainda mais essencial.

Dentre os diversos desafios recorrentes à educação na atualidade, podem-se elencar dois principais: o acesso e a qualidade. O tema deste trabalho reside no segundo desafio, ou seja, a qualidade de cursos do ensino superior no Brasil. O seu objetivo é apresentar os resultados obtidos com a aplicação da técnica estatística para a formação de Modelos de Equações Estruturais (MEE), que foi utilizada em uma investigação¹ de doutorado cujo interesse era o de montar modelos que melhor explicassem a qualidade de cursos superiores da área de gerenciamento e administração, no ano de 2006.

O MEE é uma técnica estatística que pode ser mais bem entendida como uma mistura de análise fatorial, regressão múltipla e análise de trajetória. O interesse em utilizar tal técnica está na possi-

bilidade de apreciar relações entre variáveis que se comportam como independentes, dependentes, ou ambas, em virtude de seu papel em um modelo; identificar o efeito direto, indireto e total das relações de um modelo; e utilizar diagramas de trajetória para uma maior facilidade e rapidez na visualização e entendimento de um modelo analisado.

Neste trabalho serão apresentados três modelos de equações estruturais construídos com base em construtos (variáveis não observadas) formados mediante a Análise de Classes Latentes (LCA) de dados secundários originados do Censo da Educação Superior e do Exame Nacional do Desempenho de Estudantes (ENADE).

Para tanto, o artigo está organizado em oito partes, incluindo esta, introdutória. Na segunda parte, serão apresentadas as características da investigação, ou seja, os principais conceitos trabalhados, o Modelo de Análise e o plano amostral. Na terceira parte serão apresentados os procedimentos estatísticos utilizados. A técnica de MEE será detalhada na quarta parte. A quinta parte reportará a formação do Quadro Operacional. Os resultados das aplicações da LCA e da formação dos modelos de equações estruturais serão apresentados na sexta parte. A sétima parte é reservada à conclusão e a última às referências.

Características da Investigação

A investigação da qual originou este artigo teve como objeto de estudo a qualidade de cursos superiores no Brasil, assim como os fatores que a determinam. Em razão disso, dois conceitos foram por ela abordados: o de qualidade na educação superior e o de modelos de eficácia escolar.

Independentemente da modalidade de ensino (presencial ou a distância), a qualidade da educação superior é considerada extremamente moldável à visão de mundo, ao tempo, às necessidades, à tecnologia e aos atores envolvidos em sua definição (BERTOLIN, 2007; HARVEY; GREEN, 1993; BERNETT, 1992). Considerando-se que o conceito da qualidade na educação superior é contextual, local e temporal, fica, portanto, impraticável e sem propósito eleger uma única definição. Assim, optou-se, na investigação, por identificar e

¹ Tal investigação resultou na produção da tese de doutoramento "Fatores de qualidade da educação superior: estudo sobre os dados dos cursos de administração" (CATUNDA, 2012).

trabalhar com diferentes dimensões de qualidade. Nesse sentido, as quatro dimensões de qualidade elencadas em García (2000) foram particularmente úteis a este trabalho.

A primeira dimensão apontada pela citada autora é a da **relevância**, responsável por assegurar a correspondência entre o que os estudantes aprendem e as necessidades sociais e individuais. A segunda dimensão é a da **eficácia**, na qual a qualidade seria uma relação entre os resultados logrados e os fins formulados em planos de estudos ou *standards*. A terceira dimensão reside na **eficiência**, ou seja, na relação entre os recursos empregados, os processos utilizados e os resultados obtidos, permeando assim uma visão conhecida como economicista. A quarta dimensão está vinculada à **equidade** na oferta e oportunidade de acesso e continuidade de estudos independentemente da origem e condição socioeconômica do aluno.

Por restrições metodológicas, na investigação foi abordada a dimensão eficácia, a qual relaciona os resultados logrados aos fins formulados em planos e *standards* (os alunos aprendem o que deveriam aprender). Essa escolha foi justificada por ser, na visão dos autores, a que mais representa o ideal de qualidade na educação. Esse argumento é corroborado pela reflexão de que nada adiantaria uma escola eficiente (mais egressos com menos recursos), equânime (com oportunidades a todos os alunos) ou efetiva (alinhada aos anseios da sociedade) se o aluno dela não aprendesse o que deveria aprender.

Tal dimensão de eficácia foi verificada empiricamente na investigação baseando-se em informações produzidas com o desempenho de alunos e de cursos da área de gerenciamento e administração, no Exame Nacional de Desempenho do Estudante (ENADE) de 2006. Dessa forma, levando em consideração as limitações que a restrição da escolha implica nas conclusões da investigação, foi considerada educação de qualidade (ou eficaz) aquela produzida por cursos com maiores resultados no referido exame.

Para investigar os fatores que contribuem para essa qualidade, optou-se por adotar como fundamento teórico os conhecimentos acumulados por meio das pesquisas sobre a eficácia escolar. Segundo Murillo (2003), a linha de investigação sobre

eficácia escolar compreende os estudos empíricos que buscam, por um lado, compreender a capacidade que as escolas (instituições de ensino) têm para influenciar o desenvolvimento de seus alunos e, por outro, conhecer o que torna uma escola eficaz. Para o autor, é possível distinguir dois grandes objetivos dos estudos sobre eficácia escolar: 1) Estimar a magnitude do efeito escolar e analisar suas propriedades científicas (consistência entre áreas, estabilidade, eficácia diferencial e durabilidade); e 2) Identificar os fatores de aula, escola e contexto que propiciam uma escola tornar-se eficaz.

Corroborando os estudos sobre eficácia escolar, uma gama de trabalhos busca compreender como esses fatores comportam-se em conjunto, ou seja, como eles se correlacionam entre si e como eles poderiam ser combinados para uma melhor explicação/representação do fenômeno da eficácia escolar, assim como para a identificação de modelos que proporcionem maior desenvolvimento acadêmico dos alunos (BROOKE; SOARES, 2008). Tais modelos procuram não só evidenciar “quais” fatores são importantes, mas também “como” eles influenciam a qualidade da educação.

As pesquisas em eficácia escolar, sejam elas realizadas em países desenvolvidos ou em desenvolvimento, utilizam como modelo de análise a associação entre variáveis independentes – hipoteticamente relacionadas à eficácia – e variáveis dependentes – geralmente relacionadas com o desenvolvimento cognitivo do aluno, medido por algum exame externo.

Exemplos desse tipo de pesquisa podem ser evidenciados no “Modelo Integrado da Efetividade Escolar”, desenvolvido pelo próprio Scheerens (1990), assim como no “Modelo conceitual: fatores intra e extraescolares associados ao desempenho cognitivo dos alunos”, desenvolvido no Brasil por Soares (2007).

Para a construção do Modelo Integrado da Efetividade Escolar, Scheerens (1990) analisou diferentes investigações sobre desigualdades educacionais, efeito escola, função produção escolar, escolas eficazes e efetividade docente. O autor desenhou um sistema de indicadores que informava sobre o funcionamento da escola, considerando quatro dimensões: contexto (*context*), entrada (*inputs*), processo (*process*) e produto (*outputs*). Entre os

indicadores utilizados pelo autor, estavam: 1) os fatores escolares (clima escolar seguro e ordenado; alta expectativa em relação ao rendimento dos alunos; política escolar orientada ao rendimento; liderança educativa; avaliação frequente do progresso do aluno; qualidade do currículo no que se refere ao conteúdo abrangido e à estrutura formal; e consenso e trabalho em equipe); e 2) fatores de sala de aula (tempo dedicado às tarefas; estruturação do ensino; oportunidades de aprendizagem; alta expectativa sobre o progresso dos alunos; monitoramento do progresso dos alunos; e reforço).

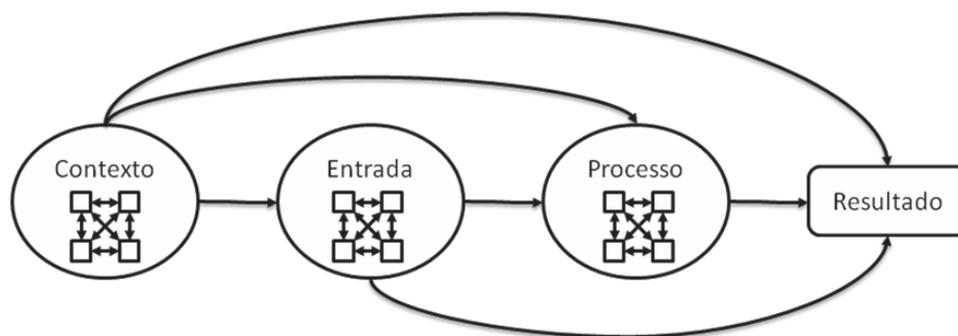
Já o modelo conceitual de Soares (2007) demonstra como fatores intra e extraescolares se associam ao desempenho cognitivo dos alunos. Na parte superior do modelo estão os fatores associados ao aluno, à família e ao grupo social de referência, os quais são considerados por alguns pesquisadores como os mais relevantes para a proficiência dos alunos. Na parte inferior estão os fatores relacionados ao contexto/sociedade e à escola. Neste último caso, os fatores são considerados importantes principalmente em situações de iniquidade da oferta e qualidade dos serviços educacionais. Para o autor, o modelo revela que são inúmeros os fatores que se associam ao desempenho dos alunos e que nenhum deles é capaz de garantir, de forma isolada, bons resultados escolares.

O modelo de análise da investigação que originou este artigo também teve como inspiração, além dos modelos citados anteriormente, o de Murillo (2008): “Modelo Analítico de Eficácia de Escolas de Primeiro Ciclo”. Tal modelo concentra-se nas variáveis de

processo e faz uma distinção clara dos fatores pertencentes à escola, à sala de aula e a ambos. Segundo o autor, os fatores relacionados à sala de aula têm maior importância sobre o rendimento do aluno do que os que estão fora da sala de aula. Nesse sentido, ganham destaque a qualidade do currículo, o clima escolar, assim como, em menor nível, o desenvolvimento profissional. Dos fatores que estão na escola, dois estão diretamente relacionados ao desempenho do aluno: sentido de comunidade (existência de normas compartilhadas) e as implicações da família. Os demais fatores estão relacionados indiretamente ao desempenho do aluno.

O modelo desta investigação, desenvolvido com base nos referenciais anteriormente citados, contemplou quatro dimensões. Três delas (Contexto, Entrada e Processo) assumiram a forma de construtos, ou seja, variáveis latentes não observadas diretamente e representadas por elipses na Figura 1, apresentada adiante. Tais construtos foram compostos baseados em variáveis que puderam ser diretamente observadas, chamadas de indicadores, representados por pequenos retângulos situados na região das elipses, e tiveram seus dados extraídos das bases do Censo da Educação Superior de 2006 e do Questionário Socioeconômico respondido pelos alunos durante o ENADE de 2006. A quarta dimensão (Resultado) – representada por um grande retângulo ao lado direito da figura – foi composta por três variáveis observáveis, relacionadas à dimensão de eficácia do conceito de qualidade, e tiveram seus dados obtidos a partir do desempenho dos cursos e dos alunos no ENADE 2006.

Figura 1 – Modelo de Análise da Investigação: Fatores de Qualidade de Cursos Superiores



Fonte: Catunda (2012, p. 82).

As partes do modelo desenvolvido neste trabalho, melhor dizendo, seus construtos e variáveis, são interconectadas por setas que indicam o sentido de suas relações. Assim, no modelo apresentado: o Resultado de um curso de ensino superior sofre influência – ou é produto – direta e indiretamente de três construtos (Contexto, Entrada e Processo); o construto Processo influencia diretamente o Resultado, mas recebe influências diretas e indiretas dos construtos Entrada e Contexto; o construto Entrada (ou insumo) influencia diretamente o Processo e direta e indiretamente o Resultado, mas recebe influências diretas do construto Contexto; e, por fim, o Contexto influencia direta e indiretamente todos os outros construtos, mas não recebe influência de nenhum deles.

Na investigação realizada, cada construto do modelo continha seu respectivo conjunto de variáveis observadas nas fontes de dados outrora citadas. Ao todo, o modelo compreendia um conjunto de 63 variáveis observadas relacionadas aos construtos, sendo 60 independentes e três dependentes.

O universo da investigação compreendeu os 2.939 cursos da área específica de gerenciamento e administração listados no Censo da Educação Superior de 2006, responsáveis por 1.035.765 alunos matriculados no primeiro semestre daquele ano (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2007a). O único critério para a formação da amostra foi o da disponibilidade de dados nos bancos acessados para as análises. Como as análises foram realizadas nos níveis aluno e curso, o número amostral variou de análise para análise, em virtude da disponibilidade de dados dos alunos e dos cursos que participaram do ENADE em 2006.

Baseado nos dados disponíveis para a pesquisa foi possível produzir dois bancos de dados: um no nível aluno e outro no nível curso. O primeiro possuía 130.270 casos (alunos), ou seja, 12,6% da população. No segundo banco de dados, dos 2.939 cursos de administração registrados em 2006, havia dados do ENADE para 1.465 deles (49,8% do total). De uma forma geral, o número amostral permitiu efetuar análises pleiteadas neste trabalho. Entretanto a grande dimensão da base de dados e a heterogeneidade das medidas das 63 variáveis

observadas impuseram aos autores a necessidade da utilização de procedimento estatístico apropriado à complexidade e à natureza categórica da maioria dessas variáveis. Tais procedimentos são descritos a seguir.

Procedimentos estatísticos

Por sua dimensão e variedade, a base de dados da investigação caracterizou-se por ser eminentemente complexa e ampla, exigindo a utilização de uma grande variedade de procedimentos estatísticos. É oportuno, para os propósitos deste trabalho, destacar a utilização das técnicas de Análise de Classes Latentes e de Modelo de Equações Estruturais, que, por sua importância para esse artigo, é abordada separadamente.

A Análise de Classes Latentes (*Latent Class Analysis* – LCA) foi utilizada para alcançar dois objetivos principais: simplificar a base de dados e formar os construtos do Modelo de Análise. A LCA é uma técnica estatística que busca a identificação do menor número de classes ou grupos de indivíduos com padrões similares de comportamento, viabilizando a criação de modelos compostos de variáveis latentes (não observáveis). Tal técnica assemelha-se à Análise Fatorial, pois ambas buscam revelar padrões de comportamento a partir de variáveis observáveis e, com isso, criar construtos que melhor representem o padrão revelado. A característica que diferencia a LCA da Análise Fatorial é a sua possibilidade de trabalhar com dados categóricos, ordinais ou nominais (COLLINS; LANZA, 2010).

Por meio da LCA foi possível identificar padrões de comportamento nas variáveis observadas e, com isso, formar os construtos (variáveis latentes) que representaram as dimensões de análise. Tais construtos serviram como variáveis independentes nos modelos de equações estruturais, necessários para compreender como os fatores de Contexto, Entrada e Processo influenciam o resultado dos alunos e cursos superiores de administração no ENADE. Cada construto construído pela LCA abrangeu apenas duas classes (padrões de resposta) de natureza qualitativa, a serem apresentadas mais adiante.

Modelos de equações estruturais

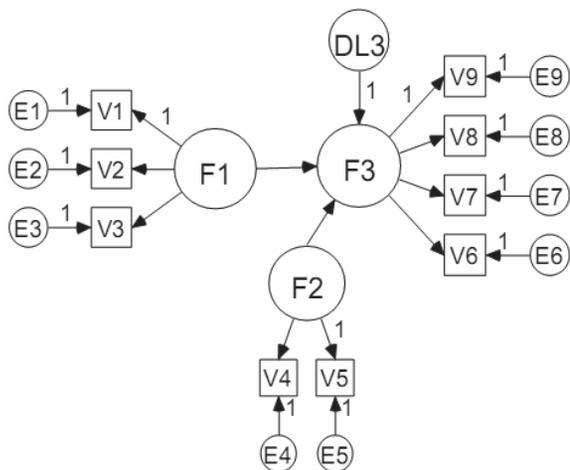
Modelos de Equações Estruturais (MEE) são um conjunto de técnicas, relativamente novas, capazes de especificar, estimar e testar relações hipotéticas entre um grupo de variáveis. Tal conjunto tem origem no trabalho seminal sobre análise de trajetórias, a qual aborda problemas de dependência entre variáveis em modelo de regressão, desenvolvido pelo geneticista Sewall Wright, em 1918, mas vai além. Por serem uma mistura de análise fatorial, regressão múltipla e análise de trajetória (PILATI, LAROS, 2007), os MEE permitem trabalhar com variáveis latentes, as quais, por serem medidas não diretamente observadas, são entendidas como fatores, construtos ou entidades hipotéticas que têm sua existência criada a partir de correlações com variáveis observáveis. Assim como é possível aos MEE apreciar relações entre variáveis observáveis e latentes, sejam elas dependentes ou independentes, ou seja, é possível analisar relações de dependência mais complexas, em que uma variável pode aparecer como dependente em uma equação e independente em outra. Dessa maneira, os MEE abrem a possibilidade para distinguir e estimar efeitos diretos, indiretos e totais entre as variáveis do modelo (CODES, 2005).

Outra característica dos MEE é o uso de diagramas de trajetória que permitem uma visualização rápida das relações de interdependência do modelo teórico a ser analisado. Como pode ser visto na Figura 2, esse diagrama é estabelecido por um conjunto de figuras geométricas conectadas por setas e dispostas em razão do papel que cada variável exerce no modelo (AMORIM et al, 2010). As variáveis nos MEE recebem as seguintes denominações, em virtude de suas características e papéis:

- Latentes ou construtos: são as variáveis que não podem ser diretamente mensuradas/observadas;
- Indicadores: são as variáveis que podem ser diretamente verificadas;
- Exógenas ou independentes: são as variáveis que não recebem influência de nenhuma variável no modelo;
- Endógenas ou dependentes: são as variáveis que recebem influência de alguma variável no modelo.

As variáveis latentes são representadas no diagrama por ciclos ou elipses, enquanto as variáveis observadas são apresentadas por quadrados ou retângulos. As variáveis são conectadas no diagrama do modelo por setas², podendo ser unidirecionais ou bidirecionais. As setas unidirecionais originam-se em uma variável exógena e se direcionam a uma variável endógena, enquanto as setas bidirecionais ligam duas variáveis endógenas. É interessante observar que uma variável pode ser tanto endógena quanto exógena se receber e enviar setas unidirecionais dentro do diagrama do modelo. Existem dois tipos básicos de modelos, os recursivos e os não-recursivos. A diferença entre os dois é que no primeiro apenas existem relações unidirecionais de determinação, enquanto no segundo, não-recursivo, há relações de dupla determinação e/ou de correlação (PILATI; LAROS, 2007). Na Figura 2, um modelo recursivo, as variáveis latentes estão codificadas com “F”, as variáveis observadas com “V”, os erros de estimação com “E”, quando forem associados a uma variável observada, e os distúrbios com “DL”, quando o erro for associado a uma variável latente.

Figura 2 – Exemplo de diagrama de MEE



Fonte: Pilati e Laros (2007, p. 208).

² Cada seta representa matematicamente uma regressão linear simples ($y = ax + b + e$); onde “a” é o coeficiente angular; “b” é coeficiente linear (intercepto); “e” o termo erro.

Segundo Codes (2005), a proposta central dos MEE é a construção de modelos que reflitam associações conceituais, tais como descritas na teoria. Nesse sentido, o procedimento metodológico consiste em elaborar e testar a plausibilidade de modelos hipotéticos que espelhem as disposições teóricas. Corroborando a citada autora, Pilati e Laros (2007) entendem que se a covariância proveniente dos dados amostrais reproduzirem os dados da população, associados ao conjunto de parâmetros impostos na construção teórica do modelo, este será considerado plausível. Para os autores, o pesquisador, ao impor parâmetros na matriz de relações entre as variáveis, antes da análise dos dados, dá ao método um caráter confirmatório. Eles ainda chamam atenção para a possível capacidade do método em produzir interpretações de causalidade no modelo final. A conclusão dos citados autores é que a causalidade entre dois fenômenos estudados em MEE está diretamente relacionada ao tipo de delineamento utilizado pelo investigador e ao seu controle metodológico – ou seja, capacidade de controlar a multiplicidade de fatores que podem, também, influenciar a variável dependente –, sendo possível que o modelo final demonstre relações corretas de casualidade.

A construção de modelos de equações estruturais envolve duas suposições e requer alguns pressupostos. Supõe-se que todas as relações causais estejam delineadas no diagrama e que estas estejam embasadas teoricamente, assim como que a natureza das relações entre as variáveis (latentes ou observadas) é linear ou pode ser linearizada por transformação (AMORIM et al, 2010). Quanto aos pressupostos, há necessidade de verificar a normalidade dos dados para a estimação quando a opção do pesquisador for o uso do método de Máxima Verossimilhança (MLE). Quando os dados não atenderem a esse pressuposto, deve-se utilizar o método Assintótico Isento de Pressupostos de Distribuição, ou ADF. Outro pressuposto é a verificação e tratamento de dados omissos ou faltosos, os quais podem ocasionar problemas de estimação dos parâmetros de um modelo estrutural (PILATI; LAROS, 2007).

Codes (2005), Pilati e Laros (2007) e Amorim et al (2010) apresentam passos metodológicos para

a construção de modelos de equações estruturais. A seguir, é apresentada uma combinação metodológica desses teóricos que atende aos propósitos deste artigo:

1. Elaboração teórica de modelos hipotéticos. Como já mencionado anteriormente, os MEE são métodos confirmatórios, não sendo adequados para explorar relações entre variáveis. Dessa maneira, o modelo hipotético a ser construído deve estar fortemente embasado em pressupostos teóricos e/ou evidências empíricas. Esse arcabouço teórico deve servir para a formulação de hipótese acerca da relação entre variáveis dentro de modelos de equações estruturais a serem testados.
2. Fundamentação teórica das medidas do modelo. É fundamental que as medidas a serem utilizadas no modelo partam do embasamento teórico anteriormente descrito e sejam válidas e fidedignas.
3. Especificação do modelo. Nessa etapa o pesquisador deve fazer uso de softwares para a estimação de modelos de equações estruturais como o AMOS, EQS, SAS ou LISREL. Para tanto, primeiro é necessário especificar o modelo de mensuração, verificando a qualidade das medidas utilizadas por meio da análise fatorial confirmatória (AFC) e, posteriormente, o modelo estrutural.
4. Identificação do modelo. Diz-se que um modelo não é identificável quando ele não consegue produzir uma solução simples para o seu sistema de equações. Para que um modelo seja identificável são necessários: a) que haja compatibilidade entre o número de parâmetros a serem estimados pelo modelo e o número de parâmetros conhecidos no modelo; b) a definição da escala de trabalho do modelo de mensuração, por meio da fixação no valor um de uma das cargas fatoriais de uma variável latente do modelo; c) que todas as medidas de erro estejam correlacionadas; d) que haja ao menos dois indicadores exclusivos para cada variável latente ou,

- quando houver apenas um indicador, que a variável latente seja mensurada sem erro; e e) que o modelo estrutural não contenha apenas variáveis observáveis.
5. Escolha do método de estimação, podendo ser: a) Quadrados Mínimos Não-Ponderados (ULS); b) Quadrados Mínimos Ponderados (GLS); c) Máxima Verossimilhança (MLE); e d) Método Assintótico Isento de Pressuposto de Distribuição (ADF), que não exige normalidade das variáveis.
 6. Avaliação do ajuste do modelo, ou seja, o quão adequadamente o modelo explica os dados. Para tanto, há uma série de critérios que podem ser utilizados. O mais tradicional deles é o Qui-quadrado, que é extremamente sensível ao tamanho das amostras. Outro índice apontado na literatura é o *Root Mean Square Error of Approximation* (RMSEA). O RMSEA compara o modelo com um modelo saturado – que estabelece todas as relações possíveis entre as variáveis –, focalizando o quão distante ou o quão desajustado o modelo se situa em relação aos dados. Um desajuste de 0,05 a 0,10 é considerado suficiente, enquanto um desajuste abaixo de 0,05 é considerado bom. Por outro lado, valores superiores a 0,10 indicam que o modelo não explica bem as correlações observadas, sendo considerados insatisfatórios. Um terceiro indicador a ser apresentado aqui é o *Root Mean Square Residual* (RMR), que indica o ajuste baseado em resíduos. Resíduos são entendidos como aspectos desconhecidos ou não levados em conta pelo modelo, ou seja, são os erros associados às variáveis latentes e observadas. Nesse sentido, quanto menor o resíduo, melhor será o modelo.
 7. Melhoria do modelo. Este estágio envolve a análise dos resultados obtidos a fim de identificar oportunidades que melhorem o ajuste do modelo. Entretanto qualquer tipo de melhoria no modelo deve estar devidamente fundamentado na teoria que embasa a construção do próprio modelo e não apenas em indicadores estatísticos.
 8. Discussão e apresentação dos resultados. Nesse estágio final o pesquisador deve discutir os resultados encontrados à luz da teoria pertinente, explicando os coeficientes apresentados pelo modelo e contribuindo com novos conhecimentos acerca do fenômeno estudado.

Quadro operacional

Conforme visto anteriormente, a base de dados da investigação continha originalmente um total de 60 variáveis (indicadores) independentes. Dessas 60 variáveis observáveis, 36 apresentaram algum padrão de comportamento verificado pela LCA e que pudesse ser associado a algum construto (variável latente ou não observada), entre as dimensões analisadas.

Como um primeiro resultado da aplicação da LCA – a simplificação da base de dados – foi possível formar, com essas 36 variáveis observadas, 11 construtos, que são apresentados no Quadro 1, o Quadro Operacional da investigação. É importante deixar claro que das três dimensões (ou construtos) originais (Contexto, Entrada e Processo) foram identificados padrões de comportamento – entre as variáveis que os compõem – que puderam originar esses 11 outros construtos (ou subconstrutos dos construtos originais). Dessa forma, do construto “Contexto” e suas variáveis (primeira e segunda colunas da Figura 2), por exemplo, puderam ser identificados os construtos “Contexto do Município” e “Contexto da Instituição de Ensino Superior (IES)” (terceira coluna do Quadro 1). Cada um desses construtos identificados após a LCA possui um padrão de comportamento entre os indivíduos (alunos ou cursos) que os distingue dos demais.

Figura 2 – Exemplo de diagrama de MEE

Construtos Antes da LCA	Variáveis observadas (Indicadores)	Construtos Após LCA
Contexto	IFDM (Desenvolvimento Municipal)	Contexto do Município
	PIB <i>per capita</i>	
	Localização (Capital ou Interior)	
	Tipo de Organização (Faculdade Ou Universidade)	Contexto da IES
	Rede de Ensino (Privada ou Pública)	
	Concorrência	
Entrada	Nº Alunos da IES	Entrada IES
	Nº Curso da IES	
	Nº Aluno Por Bibliotecas	
	Nº Aluno Por Volume	
	Nº Aluno Por Computador	
	Percentual de Docentes do Quadro	Entrada Perfil Docente
	Percentual de Docentes com 45 anos ou mais	
	Participação em Pesquisa	
	Nº Aluno Por Docente	Entrada Relação Aluno-Docente
	Nº Aluno por Docente Doutor	Entrada Perfil-Aluno
	Idade do Aluno	
	Estado Civil do Aluno	
	Nº Filhos	
	Moradia	
	Trabalho e Renda	
	Renda Familiar	Entrada SSE Aluno
Escolaridade do Pai		
Escolaridade da Mãe		
Tipo de Escola EM (Pública ou Privada)		
Processo	Frequência de Uso da Biblioteca	Processo Aluno Estudo
	Nº Livros Lidos no Ano	
	Horas Por Semana Dedicadas aos Estudos	Processo Aluno Dedicção
	Atividades Acadêmicas Desenvolvidas	Processo Perfil IES
	Iniciação Científica	
	Extensão	
	Monitoria	Processo Método IES
	Disponibilidade de Orientação Extraclasse	
	Atualização do Professor	
	Nível de Exigência	
	Integração das disciplinas	

Fonte: Pilati e Laros (2007, p. 208).

Além desses 11 construtos formados após a aplicação da LCA, foram inseridas no Modelo de Análise final as três variáveis independentes que

representavam o resultado dos alunos e dos cursos no ENADE de 2006. Tais variáveis são a seguir descritas:

- **ENADE – Conceito.** O indicador ENADE – Conceito (ou nota final do curso no ENADE) é uma composição entre as notas dos alunos ingressantes e concluintes na prova do ENADE. Tal conceito se calcula por meio da média ponderada da nota padronizada dos concluintes no componente específico, da nota padronizada dos ingressantes no componente específico e da nota padronizada em formação geral, esta última para concluintes e ingressantes (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2007b).
- **Indicador de Diferença Entre os Desempenhos Observado e Esperado (IDD).** Esse indicador de valor agregado dá uma estimativa de quanto o curso contribuiu para o desenvolvimento das habilidades acadêmicas, das competências profissionais e do conhecimento específico do aluno, levando-se em consideração o perfil do estudante que ingressou no curso. O IDD é calculado a partir da diferença entre o desempenho médio dos concluintes de um curso e o desempenho médio estimado para os concluintes desse mesmo curso. Representa, dessa forma, quanto cada curso destaca-se da média, levando em conta o perfil de seus estudantes.
- **Desempenho Geral no ENADE.** O indicador de Desempenho Geral no ENADE teve seus dados extraídos dos resultados dos alunos na prova do ENADE em 2006. Essa prova caracterizou-se por avaliar estudantes ingressantes e concluintes, abrangendo amplamente o currículo, assim como por possuir duas partes: formação geral e componente específico (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2007b).

Resultados

Com o intuito de atender ao objetivo deste artigo – apresentar os determinantes da qualidade da

educação superior –, os resultados da investigação a serem abordados nesta seção são aqueles relacionados à caracterização da amostra (após a formação dos construtos pela LCA) e, principalmente, os resultados dos Modelos de Equações Estruturais formados com base nas três variáveis dependentes e os 11 construtos, apresentados anteriormente.

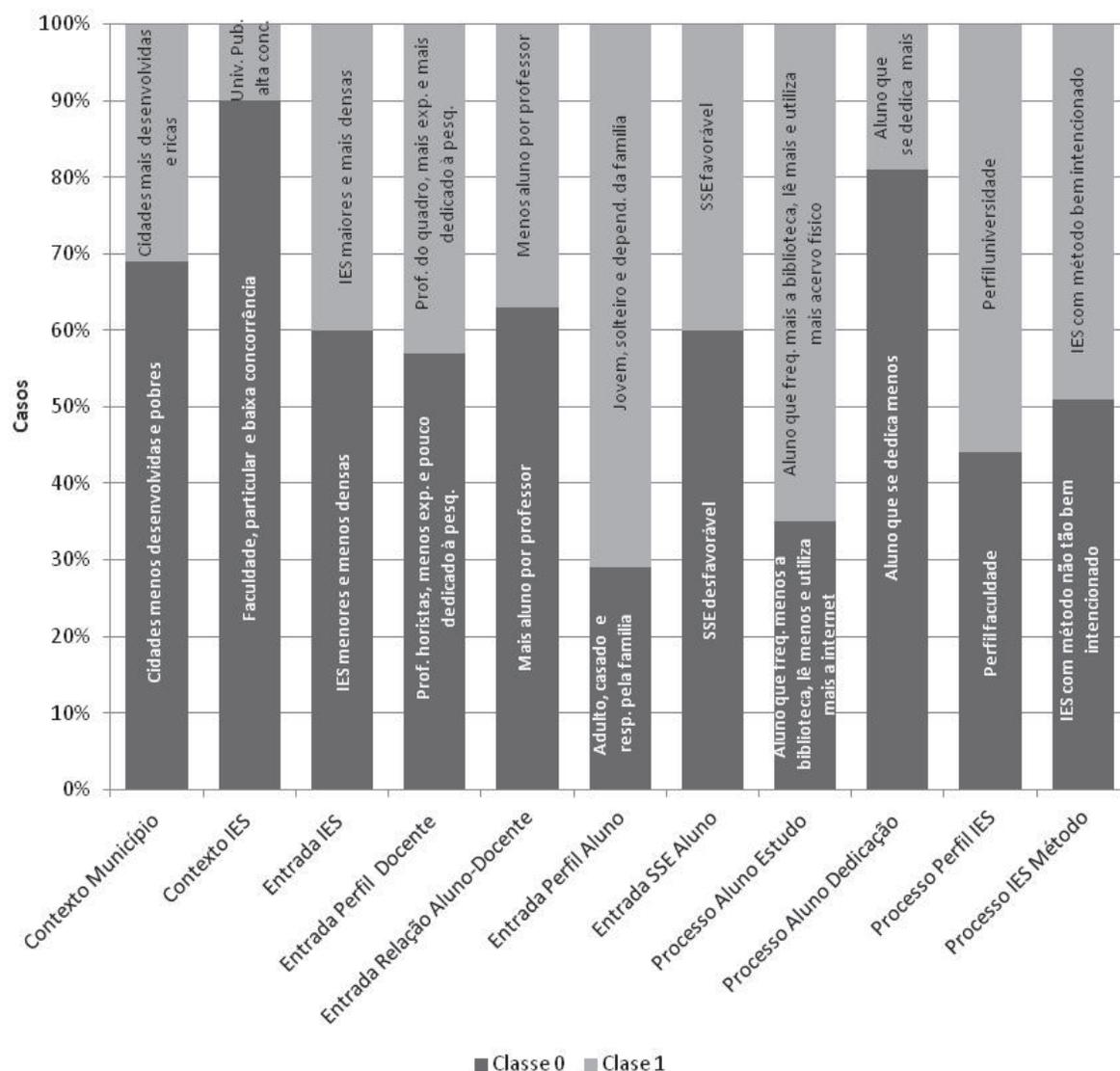
Dessa forma, a Figura 3 apresenta os 11 construtos formados e a distribuição das observações da amostra entre cada uma das duas classes de cada um dos construtos. É importante reforçar que cada classe de cada construto representa um padrão de respostas das amostras analisadas pela técnica da LCA. Nesse sentido, cada construto foi formado para que as observações fossem classificadas em um dos dois padrões de comportamento que abriga. Também é importante reforçar que cada construto relaciona-se com uma dimensão do Modelo de Análise da investigação.

O construto “Contexto Municipal”, por exemplo, pertence à dimensão “Contexto” do Modelo de Análise e possui duas classes (padrões de comportamento): Cidades menos desenvolvidas e pobres; e Cidades mais desenvolvidas e ricas. Dos cursos da área de gerenciamento e administração de 2006, analisados na investigação, aproximadamente 70% estavam em cidades menos desenvolvidas e pobres, enquanto aproximadamente 30% estavam em cidades mais desenvolvidas e ricas.

Além de simplificar a base de variáveis independentes utilizadas na investigação, a aplicação da Análise de Classes Latentes (LCA) permitiu traçar um perfil dos cursos de administração e gerenciamento em 2006, demonstrado pela Figura 3, assim como classificar cada aluno e curso observado em uma das classes dos construtos formados. Com a base de variáveis independentes, simplificada e categorizada, pode-se dar início à construção dos Modelos de Equações Estruturais, representando o Modelo de Análise da investigação e objetivando identificar a contribuição e o comportamento de cada um dos fatores determinantes da qualidade dos cursos superiores.

Antes de apresentar os modelos construídos, é importante citar que, em virtude das bases de dados possuírem dois níveis (instituição/curso e aluno), não foi possível formar um modelo que envolvesse todas as dimensões, revelando aqui

Figura 3 – Construtos e suas classes latentes



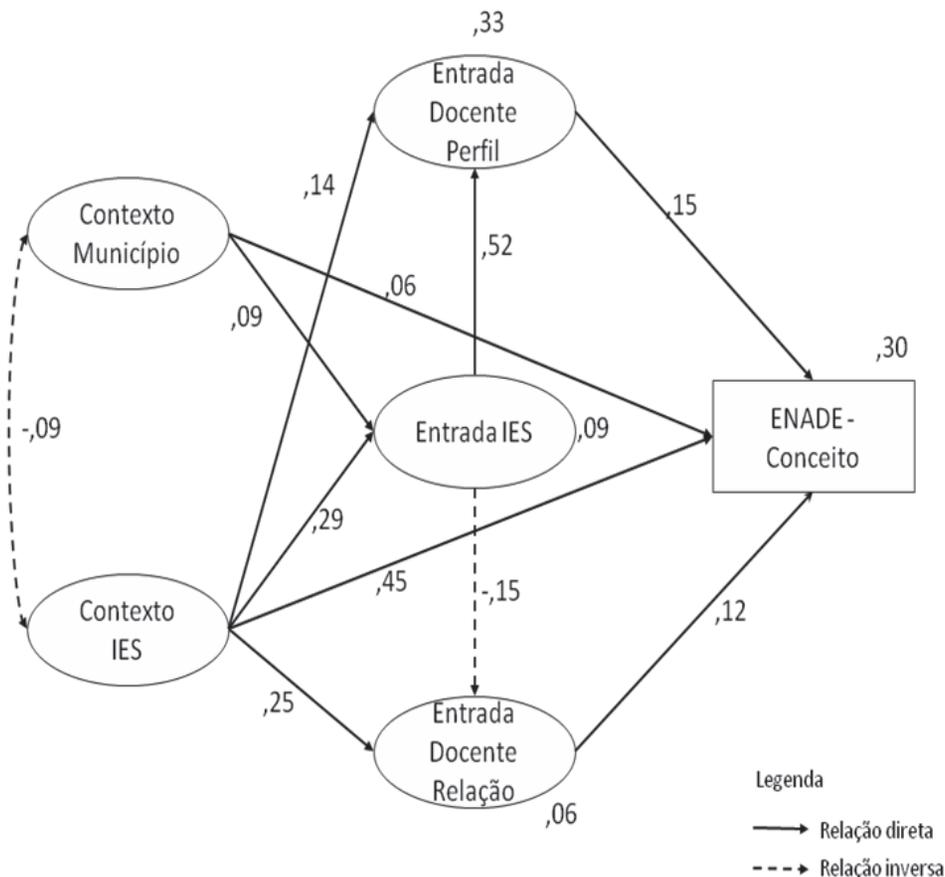
Fonte: Catunda (2012, p. 151).

uma limitação desta investigação. Com a base no nível instituição/curso foi possível construir dois modelos envolvendo as dimensões “Contexto”, “Entrada” e “Resultado”, apresentados nas Figuras 4 e 5. Assim como, com a base no nível aluno, mais um modelo foi construído envolvendo as dimensões “Entrada”, “Processo” e “Resultado”, representado na Figura 6.

No Modelo 1 (Figura 4 – Modelo 1 – Determinantes do ENADE – Conceito), todos os

construtos da base nível instituição/curso foram inseridos no diagrama, assim como a variável dependente “ENADE – Conceito”. Nesse modelo, duas variáveis assumiram função estritamente exógena (Contexto Município e Contexto IES); três assumiram função endógena e exógena (Entrada Perfil Docente, Entrada Relação Aluno-Docente e Entrada IES); e uma estritamente endógena (ENADE – Conceito).

Figura 4 – Modelo 1 – Determinantes do ENADE – Conceito



Fonte: Catunda (2012, p. 164).

De forma geral, o Modelo 1 conseguiu explicar 30% da variância do ENADE – Conceito. Apresentam-se a seguir algumas observações sobre o modelo:

- a variável (construto) que mais contribuiu para explicar o modelo é o Contexto IES;
- as setas que não apresentaram resultados significativos ($\alpha < 0,05$) foram retiradas do modelo; e
- o construto “Entrada IES” não apresentou correlação significativa com o ENADE – Conceito, porém permaneceu no modelo por ajudar a explicar dois outros construtos: “Entrada Perfil Docente” e “Entrada Relação Aluno-Docente”.

Considerando a complexidade de fatores que determinam a qualidade de um curso ou o seu resultado no ENADE, assim como a limitação de variáveis da investigação – produzida a partir de dados secundários –, o modelo apresentou um bom poder de explicação. Além disso, é possível fazer algumas interpretações interessantes com base no Modelo 1:

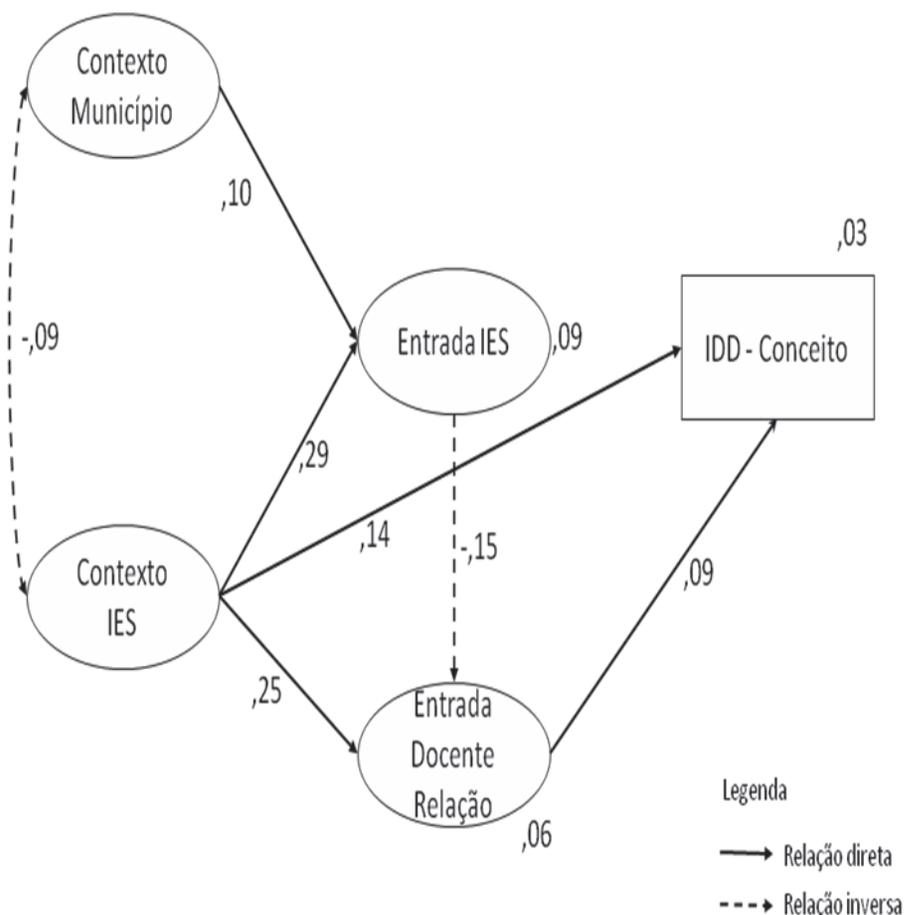
- cursos com maiores conceitos no ENADE estão mais associados a: 1) universidades públicas e com alta concorrência; 2) instituições com mais professores do quadro, experientes e mais dedicados à pesquisa; 3) instituições com menos alunos por professor e professor doutor; e, de maneira menos

- acentuada, 4) cidades mais desenvolvidas e ricas;
- não há relação significativa entre o porte e a densidade da instituição, representados pelo construto “Entrada IES”, com cursos com resultados mais altos no ENADE. Entretanto há uma associação direta entre o porte e a densidade da instituição com o perfil do professor, o que ajuda a explicar em 33% essa última variável. Ou seja, professores pertencentes ao quadro, mais experientes e com maior participação em pesquisa, estão mais associados a instituições de maior porte e densidade;
 - o porte e a densidade também estão associados à quantidade de alunos por professor

e professor doutor, entretanto de maneira inversa. Isso faz sentido na medida em que o valor menor na regressão refere-se às instituições com maior número de alunos por professor e professor doutor. A contribuição da Entrada IES, assim como do Contexto da IES, ajuda a explicar apenas 6% dessa variável; e

- o porte e a densidade são pouco explicados, apenas 9%, sendo o Contexto da IES a variável (construto) que mais ajuda na sua explicação. Ou seja, IES grandes e densas estão mais associadas às universidades públicas e com alta concorrência, assim como aos municípios mais desenvolvidos e ricos.

Figura 5 – Modelo 2 – Determinantes do IDD – Conceito



Fonte: Catunda (2012, p. 167).

No Modelo 2 (Figura 5 – Modelo 2 – Determinantes do IDD – Conceito) todos os construtos da base nível instituição/curso foram inseridos no diagrama, assim como a variável dependente “IDD – Conceito”. Da mesma forma como ocorreu com o Modelo 1, duas variáveis assumiram função estritamente exógena (Contexto Município e Contexto IES); três assumiram função endógena e exógena (Entrada Perfil Docente, Entrada Relação Aluno-Docente e Entrada IES); e uma estritamente endógena (IDD – Conceito).

De forma geral, o Modelo 2 conseguiu estimar apenas 3% da variância do IDD – Conceito. Apresentam-se a seguir outras observações sobre o modelo:

- a variável (construto) que mais contribui para explicar o modelo também foi o Contexto IES;
- as setas que não apresentaram resultados significativos ($\alpha < 0,05$) foram retiradas do modelo;
- o construto “Entrada Perfil Docente” saiu do modelo por não apresentar associação significativa com o IDD – Conceito e nem com nenhum outro construto; e
- o construto “Entrada IES” não apresentou correlação significativa com o IDD – Conceito, porém permaneceu no modelo por ajudar a explicar o construto “Entrada Relação Aluno-Docente”.

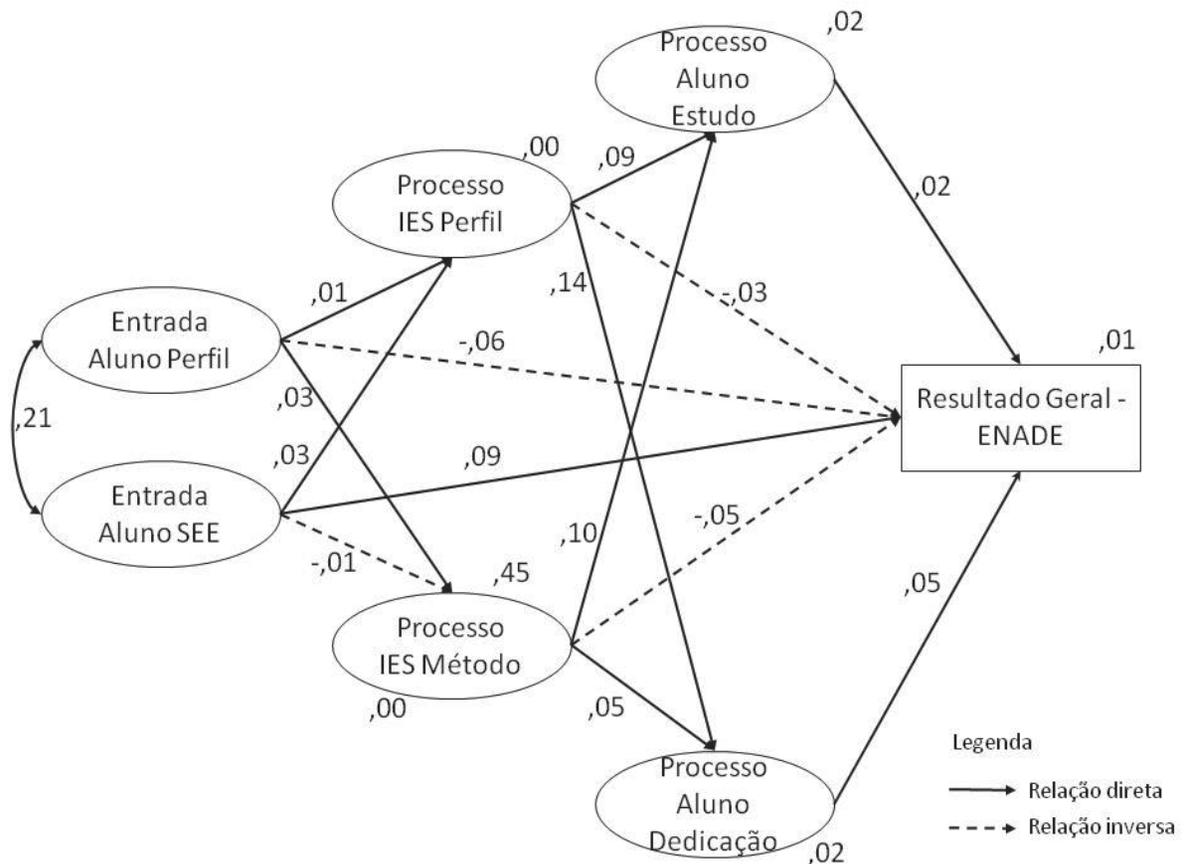
Considera-se que o Modelo 2 tem pouco poder de explicar a variável dependente IDD – Conceito. Essa conclusão leva à ideia que possivelmente as variáveis analisadas por essa pesquisa, originadas em sua maioria no Censo da Educação Superior e no Questionário Socioeconômico do ENADE, não foram úteis para identificar os determinantes do IDD – Conceito. Esse resultado é corroborado pelos achados de Bartholomeu (2011, p. 130), os quais indicam que a variância de desempenho dos cursos no IDD é mais bem explicada pelo desempenho dos alunos no ENEM, “restando pouca variância de desempenho final para ser explicado pelo que o curso agrega ao sujeito”. A partir de variáveis intraescolares, Soares, Ribeiro e Castro (2001) conseguiram explicar no máximo 6% da variância

do IDD de cursos de direito, administração e engenharia civil, como visto no Marco Teórico dessa pesquisa. A despeito disso, é possível fazer algumas interpretações interessantes baseadas no Modelo 2:

- cursos com maiores conceitos no IDD estão mais associados a: 1) universidades públicas e com alta concorrência; e 2) instituições com menos alunos por professor e professor doutor;
- não há relação significativa entre o porte e a densidade da instituição, representados pelo construto Entrada IES, com cursos com resultados mais altos no IDD. Entretanto o porte e a densidade estão inversamente associados à quantidade de alunos por professor e professor doutor. Isso faz sentido na medida em que o valor menor na regressão refere-se às instituições com maior número de alunos por professor e professor doutor. A contribuição do porte e da densidade, assim como do contexto da IES, ajuda a explicar apenas 6% dessa variável; e
- não há relação significativa entre o contexto do município e cursos com resultados mais altos no IDD. Entretanto o contexto do município ajuda a explicar o porte e a densidade da IES. Ainda assim, o porte e a densidade são pouco explicados, apenas 9%, sendo o contexto da IES a variável que mais a ajuda na sua explicação. Ou seja, IES grandes e densas estão mais associadas às universidades públicas e com alta concorrência, assim como aos municípios mais desenvolvidos e ricos.

De forma geral, o Modelo 3, apresentado na Figura 6, conseguiu estimar apenas 1% da variância do Resultado Geral do ENADE. Além disso, suas relações internas pouco serviram para explicar as variáveis endógenas do próprio modelo. Tal fato pode ter decorrido da natureza da fonte de dados (questionário opinativo respondido pelos alunos) ou pelo tipo de informação das variáveis, as quais apresentaram baixa associação nas análises bivariadas. Sendo assim, considerando os conceitos e a metodologia aplicada nessa investigação, tal modelo não se mostrou interessante para explicar a qualidade dos cursos superiores.

Figura 6 – Modelo 3 – Determinantes do Resultado Geral no ENADE



Fonte: Catunda (2012, p. 171).

Conclusão

Este artigo teve o propósito de apresentar os resultados obtidos a partir da aplicação da técnica estatística para a formação de Modelos de Equações Estruturais (MEE), utilizada para montar modelos que melhor explicassem a qualidade de cursos superiores. Para tanto, fez uso dos resultados de uma investigação com dados secundários de alunos e de cursos da área de gerenciamento e administração, assim como de seus resultados no ENADE de 2006. Ao todo, foram elaborados três modelos – um para cada uma das variáveis dependentes relacionadas ao conceito de qualidade da educação superior.

Os modelos ora montados apresentaram diferentes capacidades de explicação das variáveis dependentes. O Modelo 1 conseguiu explicar 30%

da variável **ENADE – Conceito**, enquanto os Modelos 2 e 3 conseguiram explicar, respectivamente, 3% do **IDD – Conceito** e 1% do **Resultado Geral no ENADE**. Levando-se em conta estritamente o primeiro modelo, o uso das técnicas de LCA e MEE permitiu identificar um perfil de instituição de ensino superior que alcançou maiores conceitos no ENADE de 2006: universidades públicas e com alta concorrência; com mais professores do quadro, experientes e mais dedicados à pesquisa; com menos alunos por professor e professor doutor; e localizadas em cidades mais desenvolvidas e ricas.

A despeito dos Modelos 2 e 3 não terem produzido melhores resultados, de uma forma geral, os três modelos apresentados trazem contribuições às discussões acerca de políticas educacionais na medida em que acrescentam novo conhecimento a

temas conflituosos, tais como: operacionalização da educação (público x privado); dimensão e infraestrutura das instituições; perfil da docência; perfil e atitudes dos alunos; e método de ensino.

É importante registrar que os resultados devem ser compreendidos levando em consideração as limitações da investigação, entre elas: utilização exclusiva de dados secundários produzidos pelo Ministério da Educação e pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP); inviabilidade da formação de modelos que envolvessem todas as dimensões analisadas (contexto, entrada, processo e resultados); a não utilização de modelos multiníveis; a restrição do conceito de qualidade à dimensão de eficácia; a não utilização de recorte temporal longitudinal;

e a limitação da amostra aos cursos da área de gerenciamento e administração.

Os resultados aqui apresentados, seja pela escassez de pesquisa com dados da educação superior, pela aplicação de técnica estatística recentemente desenvolvida (LCA) ou ainda pouco utilizada (MEE), assim como pelas limitações da própria investigação, devem ser interpretados como uma fase inicial na compreensão dos fatores que determinam a qualidade da educação superior. Nesse sentido, novas investigações devem ser realizadas, seja para confirmar os modelos aqui reportados em outras áreas da educação superior, seja para aprofundar e aperfeiçoar a utilização da técnica ou para explorar novas dimensões e indicadores desse fenômeno (qualidade), eminentemente complexo.

REFERÊNCIAS

- AMORIM, Leila D. A. F. et al. Structural equation modeling in epidemiology. *Cadernos de Saúde Pública*, Rio de Janeiro, v. 26, n. 12, p. 2251-2262, dez. 2010.
- BARTHOLOMEU, Daniel. **O uso do ENEM e ENADE em um delineamento longitudinal para avaliação da qualidade dos cursos do ensino superior**. 2011. 170 f. Tese (Doutorado em Psicologia) – Universidade São Francisco, Itatiba, SP, 2011.
- BERNETT, R. **Improving higher education – total quality care**. Bristol, USA: SRHE and Open University Press, 1992.
- BERTOLIN, J. C. G. **Avaliação da qualidade do sistema de educação superior brasileiro em tempos de mercantilização: período de 1994-2003**. 2007. 282 f. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.
- BROOKE, N.; SOARES, J. F. (Org.). **Pesquisa em eficácia escolar: origens e trajetórias**. Belo Horizonte: Editora UFMG, 2008.
- CATUNDA, Arturo Cavalcanti. **Fatores de qualidade da educação superior: estudo sobre os dados dos cursos de administração**. 2012. 231 f. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal da Bahia, Salvador, 2012.
- CODES, Ana Luiza Machado de. **Modelagem de equações estruturais: uma contribuição metodológica para o estudo da pobreza**. 2005. 297 f. Tese (Doutorado em Ciências Sociais) – Programa de Pós-Graduação em Ciências Sociais, Universidade Federal da Bahia, Salvador, 2005.
- COLLINS, L. M.; LANZA, S. T. **Latent class and latent transition analysis**. New Jersey: Wiley, 2010.
- GARCÍA, M. G. Evaluación y Calidad de los Sistemas Educativos. In: RAMÍREZ, Teresa G. (Org.). **Evaluación y Gestión de la Calidad Educativa**. Málaga: Ediciones Aljibe, 2000.
- HARVEY, L.; GREEN, D. Defining Quality. *Assessment & Evaluation in Higher Education*, v. 18, n. 1, p. 9-26, abr. 1993.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Sinopse da educação superior 2006**. Planilha Excel. Brasília, 2007a. Disponível em: <<http://www.inep.gov.br/superior/censosuperior/sinopse/>>. Acesso em: 02 jan. 2010.
- _____. **ENADE 2006. Relatório síntese: curso de Administração**. Brasília, 2007b.

MURILLO, F. J. Una panorámica de la investigación iberoamericana sobre eficacia escolar. **REICE - Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**, v. 1, n. 1, p. 1-14, 2003.

_____. Hacia un modelo de eficacia escolar. Estudio multinivel sobre los factores de eficacia de las escuelas españolas. **REICE - Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**, v. 6, n. 1, p. 4-28, 2008.

PILATI, Ronaldo; LAROS, Jacob Arie. Modelos de equações estruturais em Psicologia: conceito e aplicações. **Psicologia: Teoria e Pesquisa**, Brasília, v. 23, n. 2, p. 205-216, abr./jun. 2007.

SCHEERENS, J. School effectiveness research and the development of process indicators of school functioning. **School Effectiveness and School Improvement**, v. 1, n. 1, p. 61-80, 1990.

SOARES, J. F. Melhoria do desempenho cognitivo dos alunos do ensino fundamental. **Cadernos de Pesquisa**, São Paulo, v. 37, n. 130, p. 135-160, abr. 2007. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-15742007000100007&lng=en&nrm=iso>. Acesso em: 28 mar. 2010.

SOARES, J. F.; RIBEIRO, L.; CASTRO, C. M. Valor agregado de instituições de ensino superior em Minas Gerais para os cursos de Direito, Administração e Engenharia Civil. **DADOS**, Rio de Janeiro, v. 44, n. 2, p. 363-396, 2001. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0011-52582001000200005&lng=en&nrm=iso>. Acesso em: 18 jan. 2012.

Recebido em 27.06.12

Aprovado em 03.08.12