

A auto-avaliação de IES: um modelo para a avaliação das disciplinas curriculares segundo a percepção do corpo discente

ANDRÉ LUÍS POLICANI FREITAS
GUDELIA MORALES DE ARICA

LEPROD, Centro de Ciencia e Tecnologia,
Universidade Estadual do Norte Fluminense, Brasil

1. Introdução

Nos últimos anos, o Censo da Educação Superior no Brasil revela que o sistema brasileiro de educação superior vivencia um processo de expansão acelerado. Esta expansão traduz-se no aumento do número de Instituições de Ensino Superior (IES) e de cursos (presenciais e semi-presenciais) e, por conseguinte, no aumento do número de vagas oferecidas e do número de professores e funcionários destas instituições.

Apenas para exemplificação: segundo Oliveira *et al.* (2005), "a partir de 1998 houve um crescimento vertiginoso do número de cursos de Engenharia de Produção no Brasil, saltando dos 38 registrados em 1997 para, aproximadamente, 200 cursos em 2005, registrando-se a criação de cerca de 20 cursos por ano". Segundo o INEP (2006), em 1997 o percentual de cursos de Engenharia de Produção era de 47% em IES públicas e 53% em IES privadas, passando para 71% em IES privadas e 29% em IES públicas, em 2005.

Com o intuito de avaliar e monitorar o desempenho das Instituições de Ensino Superior, dos cursos e dos estudantes, foi criado o SINAES - Sistema Nacional de Avaliação da Educação Superior (Lei n.º 10.861, de 14 de abril de 2004). O SINAES propõe-se avaliar todos os aspectos no âmbito acadêmico, dentre os quais citam-se: o ensino, a pesquisa, a extensão, a responsabilidade social, o desempenho dos alunos, a gestão da instituição, o corpo docente, as instalações (com ênfase nas bibliotecas). Em particular, o SINAES integra três modalidades principais de instrumentos de avaliação, aplicadas em diferentes momentos:

- 1) Avaliação das Instituições de Educação Superior (AVALIES) – é o centro de referência e articulação do sistema de avaliação que se desenvolve em duas etapas:
 - a) Auto-avaliação – coordenada pela Comissão Própria de Avaliação (CPA) de cada IES, a partir de 1.º de setembro de 2004;
 - b) Avaliação externa – realizada por comissões designadas pelo INEP, segundo diretrizes estabelecidas pela CONAES.

Revista Iberoamericana de Educación

ISSN: 1681-5653

n.º 44/7 – 10 de enero de 2008

EDITA: Organización de Estados Iberoamericanos
para la Educación, la Ciencia y la Cultura (OEI)



- 2) Avaliação dos Cursos de Graduação (ACG) – avalia os cursos de graduação por meio de instrumentos e procedimentos que incluem visitas *in loco* de comissões externas. A periodicidade desta avaliação depende diretamente do processo de reconhecimento e de renovação de reconhecimento a que os cursos estão sujeitos.
- 3) Avaliação do Desempenho dos Estudantes (ENADE) – aplica-se aos estudantes do final do primeiro e do último ano do curso, estando prevista a utilização de procedimentos amostrais.

Neste contexto, os Centros Universitários e as Universidades são motivadas a desenvolver um mecanismo ou procedimento de auto-avaliação que contemple as atividades de ensino, de pesquisa e de extensão. Segundo o INEP (2002), os aspectos a serem mensurados por este procedimento seriam: a existência de órgão ou comissão permanente de avaliação, abrangência do projeto de auto-avaliação, participação da comunidade acadêmica nos processos de auto-avaliação, divulgação dos resultados da auto-avaliação, ações acadêmico-administrativas em função da auto-avaliação, articulação entre a interpretação dos resultados das avaliações realizadas pelo MEC, das avaliações realizadas por outros agentes externos e os da auto-avaliação da IES.

Apesar da importante contribuição destas modalidades do SINAES no âmbito da avaliação do Ensino Superior, em geral, estas não consideram o ponto de vista dos dois elementos principais e essenciais em um processo de formação educacional: o corpo docente e os alunos. Adicionalmente, observa-se que grande parte dos critérios (itens) de avaliação utilizados possui natureza tipicamente técnica, não abordando a questão da Qualidade do Ensino em sua essência prática.

A modo de visualização, no contexto da avaliação do corpo docente, alguns dos critérios usualmente empregados são: a qualificação, a experiência profissional e a carga horária dedicada à Instituição de Ensino. De fato estes critérios são muito importantes, mas não necessariamente asseguram um ensino de qualidade. Entretanto, critérios não menos importantes para a qualidade do Ensino, tais como a didática utilizada em sala de aula (ou laboratório), a presteza para proporcionar o esclarecimento de dúvidas e a pontualidade em nenhum momento são precisamente considerados pelos programas de avaliação das Instituições de Ensino Superior até então desenvolvidos.

Dentre outros aspectos, o desempenho inadequado da IES segundo os três critérios supracitados pode se refletir diretamente no grau de absorção do conhecimento adquirido pelos alunos, na adequação do conteúdo programático das disciplinas a um determinado curso de nível superior (Graduação) e no cumprimento do conteúdo programático das disciplinas. As conseqüências finais podem ser indesejáveis, tais como: o aumento dos índices de reprovação, o aumento da taxa de evasão escolar, resultados insatisfatórios no Exame Nacional de Cursos, a redução do número de concluintes e ingressantes, etc. Particularmente, estas conseqüências são alguns dos índices utilizados pelo INEP para monitorar o desenvolvimento do sistema brasileiro de educação superior.

Neste contexto, verifica-se que ainda existe uma lacuna no que diz respeito à questão da avaliação da qualidade do Ensino Superior. Em especial, observa-se que ainda existe uma indefinição dos organismos governamentais direcionados a respeito de uma abordagem metodológica para o tratamento do problema em questão, problema este que requer avaliações mais detalhistas, portanto específicas da CPA de cada IES e não das comissões externas alocadas pelo INEP pela dificuldade logística de implementação.

Além disso, acredita-se que as informações fornecidas pelos docentes e pelos alunos da IES podem contribuir positiva e preventivamente para a melhoria da qualidade do ensino.

Com o intuito de contribuir para o preenchimento desta lacuna e de dar continuidade às discussões e pesquisas direcionadas no sentido do aprimoramento contínuo do sistema de avaliação do ensino superior, este trabalho apresenta um instrumento de auto-avaliação desenvolvido no âmbito da avaliação de disciplinas curriculares, segundo a percepção do corpo discente. Para verificar a confiabilidade do instrumento proposto, utilizou-se um dos procedimentos estatísticos mais utilizados para este fim – o coeficiente alfa de Cronbach.

Sucintamente, este artigo está estruturado da seguinte forma: a seção 2 apresenta uma breve descrição da Teoria Clássica da Medição e a sua relação com a mensuração da confiabilidade dos instrumentos de pesquisa; a seção 3 descreve o procedimento para o cálculo do coeficiente alfa de Cronbach; a seção 4 apresenta um experimento destinado a investigar a aplicação do instrumento proposto para a avaliação das disciplinas curriculares do curso de Graduação em Engenharia de Produção de uma universidade pública estadual; as seções 5 e 6 apresentam, respectivamente, as considerações e restrições à implementação do modelo proposto; a seção 7 apresenta as conclusões deste trabalho; e, finalmente, nos Apêndices são apresentados o modelo de questionário utilizado para a avaliação das disciplinas, o modelo de Relatório com o resultado da disciplina avaliada no experimento realizado, além de uma breve descrição dos Coeficientes de Correlação utilizados neste artigo.

2. A confiabilidade dos instrumentos de medição

Em toda pesquisa de opinião, a captação de informações dos avaliadores é uma das etapas de maior importância no processo de avaliação. Esta etapa deve ser realizada através do emprego de um instrumento de medição eficaz e preciso, pois, caso contrário, as informações obtidas podem não ser representativas da percepção dos avaliadores¹. Dentre os instrumentos de pesquisa tradicionalmente utilizados, destaca-se o formulário de pesquisa (tradicionalmente denominado questionário) por sua praticidade de aplicação (pode ser entregue pessoalmente ao avaliador, deixado ao alcance dos mesmos, enviados pelo correio, por *e-mail*, etc.) e também pelo seu baixo custo. Entretanto, eventualmente pode ocorrer um baixo índice de retorno de questionários respondidos e um alto índice de itens não respondidos.

De maneira sucinta, a Teoria Clássica da Medição considera que em toda medição o valor observado X é composto aditivamente por duas variáveis: o valor verdadeiro (T) e erro aleatório de medição (E). Segundo Hayes (1995), a variância de cada conjunto de variáveis está associada de forma semelhante às próprias variáveis, ou seja:

$$Var(X) = Var(T) + Var(E) \quad (1)$$

Assim, à medida que a variância associada a erros aleatórios diminui, a variância dos valores observados se aproxima da variância dos valores verdadeiros, o que representa maior confiabilidade nas medições (e conseqüentemente maior confiabilidade atribuída ao instrumento de coleta de dados utilizado).

¹ Em pesquisas de *Marketing*, os avaliadores em geral são os consumidores de determinado produto ou serviço. No contexto deste artigo, consideram-se avaliadores os alunos que cursam cada disciplina.

Conceitualmente, a confiabilidade reflete a intensidade com que os valores observados estão correlacionados aos valores verdadeiros. Ou ainda, segundo Hayes (1995), “a confiabilidade é definida como o grau com que as medições estão isentas de erros aleatórios”.

Entretanto, como em muitos problemas reais os valores verdadeiros das variáveis não são conhecidos, é impossível determinar a confiabilidade, diretamente, através do cálculo da correlação entre os valores observados e os verdadeiros. Segundo Trochim (2003) existem quatro procedimentos para estimar a confiabilidade de medições:

- a) Confiabilidade entre avaliadores, que é utilizada para obter o grau com que diferentes avaliadores fornecem estimativas consistentes sobre o mesmo fenômeno.
- b) Confiabilidade de teste/reteste, utilizada para estimar a consistência de uma medida em dois momentos distintos.
- c) Confiabilidade paralela, utilizada para determinar a consistência dos resultados de duas pesquisas distintas elaboradas com o mesmo objetivo.
- d) Confiabilidade de consistência interna, utilizada para determinar a consistência dos resultados da avaliação de itens de uma pesquisa com o resultado geral da pesquisa, o que representa uma mensuração da confiabilidade desse resultado.

Particularmente, este artigo concentra-se na investigação da confiabilidade de consistência interna do instrumento de avaliação proposto através do emprego de um dos procedimentos estatísticos mais utilizados para a mensuração da consistência interna - o coeficiente alfa de Cronbach, brevemente descrito na seção 2.1. Adicionalmente, utilizam-se coeficientes de correlação (Correlação Inter-Item e Correlação Item-Total) para a obtenção da correlação existente entre os itens que compõem o questionário.

2.1. O coeficiente Alfa de Cronbach

Apresentado por Lee J. Cronbach em 1951, o coeficiente α de Cronbach é uma das estimativas da confiabilidade de um questionário que se aplica em uma pesquisa. Dado que todos os itens de um questionário utilizam a mesma escala de medição, o coeficiente, α , com $\alpha \in [0,1]$, é calculado a partir da seguinte equação:

$$\alpha = \left(\frac{k}{k-1} \right) \left(1 - \frac{\sum_{i=1}^k S_i^2}{S_t^2} \right) \quad (2)$$

onde k é o número de itens do questionário, S_i^2 é a variância do item i e S_t^2 é a variância total do questionário.

Assim, o coeficiente α pode ser calculado em duas etapas: primeiramente é necessário calcular a variância de cada coluna i , denotada por S_i^2 e, em seguida somar todas estas variâncias. Na segunda etapa deve-se obter a soma total dos julgamentos de cada avaliador (elementos da última coluna da tabela 1 e, em seguida calcular a variância destas somas.

TABELA 1

Matriz de julgamentos dos n avaliadores à luz dos k itens do questionário (Cronbach, 2004)

AVALIADORES	ITENS						SOMA OU TOTAL
	1	2	...	i	...	k	
1	X_{11}	X_{12}	...	X_{1i}	...	X_{1k}	X_1
2	X_{21}	X_{22}	...	X_{2i}	...	X_{2k}	X_2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
p	X_{p1}	X_{p2}	...	X_{pi}	...	X_{pk}	X_p
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
n	X_{n1}	X_{n2}	...	X_{ni}	...	X_{nk}	X_n
				S_i^2			S_t^2

Considerando que a variância total pode ser reestruturada como sendo a soma da variância dos itens adicionada a duas vezes a covariância destes itens, a equação 2 pode ser reescrita de uma forma alternativa que, segundo Peter (1979), facilita os cálculos computacionais:

$$\alpha = \left(\frac{k}{k-1} \right) \left(1 - \frac{\sum_{i=1}^k \sigma_i^2}{\underbrace{\sum_{i=1}^k \sigma_i^2}_A + 2 \underbrace{\sum_{i > j}^k \sigma_{ij}}_B} \right), \quad (3)$$

Na equação 3, o termo A representa a soma das variâncias dos itens (a soma dos elementos da diagonal principal da matriz de covariância). O termo B representa duas vezes a soma da covariância dos elementos, ou seja, duas vezes a soma dos elementos não pertencentes à diagonal principal da matriz de covariância. Para uma melhor compreensão, a tabela 2 ilustra os elementos da matriz de covariância.

TABELA 2

Matriz de Covariância para um questionário hipotético de seis itens

MATRIZ DE COVARIÂNCIA	ITENS					
	1	2	3	4	5	6
1	Cov_{11}	Cov_{12}	Cov_{13}	Cov_{14}	Cov_{15}	Cov_{16}
2		Cov_{22}	Cov_{23}	Cov_{24}	Cov_{25}	Cov_{26}
3			Cov_{33}	Cov_{34}	Cov_{35}	Cov_{36}
4				Cov_{44}	Cov_{45}	Cov_{46}
5					Cov_{55}	Cov_{56}
6						Cov_{66}

Entretanto, segundo a literatura científica existente, diversos fatores que podem influenciar na confiabilidade de questionários, dentre os quais se citam:

- a) NÚMERO DE ITENS: a confiabilidade do questionário pode ser aumentada quando mais itens são incluídos, o que seria equivalente ao conceito de diminuir o erro de amostragem, através do aumento do tamanho da amostra (Hayes,1995). Entretanto, os itens adicionais devem ser representativos do mesmo conceito que está sendo medido e o resultado da avaliação de um item não deve interferir na avaliação do item subsequente. Além disso, o número de itens não deve ser excessivo, pois este fato pode ocasionar respostas impulsivas e relapsas, além de poder aumentar a incidência de itens sem resposta, devido à fadiga/desinteresse do avaliador.
- b) TEMPO DE APLICAÇÃO DO QUESTIONÁRIO: a aplicação do questionário, segundo um período pré-definido também pode contribuir para a ocorrência de respostas impulsivas e relapsas, além de aumentar a incidência de itens sem avaliação. Segundo Cronbach (2004), “quando o questionário é muito grande, o alfa não deve ser utilizado. Por exemplo, em um teste aplicado por um tempo limitado, um apreciável número de estudantes parou de responder antes de chegar aos últimos itens”.
- c) A AMOSTRA DE AVALIADORES: uma amostra de pessoas semelhantes pode resultar em um questionário de baixa confiabilidade. Por exemplo, se na avaliação de um item todas as pessoas da amostra assinalarem um valor verdadeiro 2 (considerando uma escala de valores absolutos que varia de -2 a 2), não haverá variância neste item e, por conseguinte, a confiabilidade calculada terá valor zero. Assim, segundo Hayes (1995), “se desejarmos obter medidas com alta confiabilidade, precisaremos basear essas medidas em uma amostra de pessoas que sejam heterogêneas no que diz respeito ao conceito que está sendo medido”.

Entretanto, é importante destacar que as pessoas geralmente possuem percepções e interpretações distintas sobre os itens em avaliação. Desta forma, considerando uma amostra significativa de avaliadores, é muito improvável que haja uma concordância absoluta a respeito do valor de determinado item. Porém, quando os avaliadores são semelhantes em sua natureza e formação profissional (*experts*), possivelmente haverá menor variabilidade nos julgamentos, o que pode conduzir a menores valores de confiabilidade do questionário.

4. Modelo para avaliação de disciplinas segundo a percepção dos alunos

Com o intuito de contribuir para o sistema de auto-avaliação de Instituições de Ensino Superior (IES) recomendado pelo SINAES/INEP, o Laboratório de Engenharia de Produção da Universidade Estadual do Norte Fluminense (LEPROD/UENF) vem dedicando-se ao desenvolvimento de um instrumento de pesquisa capaz de captar as percepções dos alunos quanto ao desempenho das disciplinas que compõem o currículo do Curso de Graduação em Engenharia de Produção.

4.1. Estrutura do modelo de avaliação (questionário) atualmente utilizado

Até o momento foram realizadas três pesquisas de avaliação, sendo estas realizadas sempre nas últimas semanas letivas dos últimos três semestres. O modelo inicial do questionário de avaliação das disciplinas, assim como os modelos aperfeiçoados, foram previamente submetidos à análise dos membros pertencentes ao Colegiado do Curso de Graduação em Engenharia de Produção da UENF.

O modelo de questionário atualmente está estruturado da seguinte forma (no Apêndice 1 apresenta-se o modelo de questionário utilizado):

- 1) IDENTIFICAÇÃO: nesta etapa é necessário identificar o nome da disciplina a ser avaliada e o professor responsável por esta. Além disso, o aluno-avaliador deve informar opcionalmente o seu nome, o período cursado na IES, o seu coeficiente de rendimento (CR) acumulado e a natureza do Ensino Médio (Escola Estadual, Federal ou Particular).
- 2) DIMENSÕES E ITENS: cada aluno deverá estabelecer sua percepção quanto ao Grau de Desempenho da disciplina à luz de Dimensões, cada qual composta por Itens a ela correlacionados. Tais Dimensões e Itens buscam retratar os elementos presentes e intrínsecos no processo de ensino-aprendizagem comum à maioria das disciplinas, sendo assim denominadas: Disciplina (conteúdo), Disciplina (provas e testes), Material Didático, Professor da disciplina e Auto-avaliação. Em especial, a Dimensão "Auto-avaliação" propõe fazer com que o aluno reflita e se posicione quanto a sua participação no processo de ensino-aprendizagem, visto que ele representa um elemento essencial neste processo. Nesta etapa de avaliação do desempenho, recomenda-se o uso da escala tipo Likert, cujos conceitos associados ao desempenho variam de "Muito Ruim" a "Muito Bom". Segundo Mattar (1996), o uso desta escala faz com que os avaliadores sejam solicitados não somente a informar se o desempenho do objeto avaliado é Bom ou Ruim, mas também informar a intensidade deste desempenho.

Adicionalmente, considera-se a Dimensão "Avaliação Complementar". Nesta etapa, o aluno deverá estabelecer o Grau de Concordância à luz de alguns itens, através do uso da escala tipo Likert, cujos conceitos variam de "Discordo Totalmente" a "Concordo Totalmente". Tais itens apresentam conceituação associada a um ou mais itens pertencentes às Dimensões anteriormente avaliadas, sendo possível verificar a coerência das respostas.

- 3) INFORMAÇÕES ADICIONAIS: esta etapa do questionário consiste em um "espaço aberto" para que cada aluno possa registrar por escrito seus comentários a respeito do conteúdo da disciplina e do material didático e também a respeito do professor da disciplina, além de suas sugestões para melhorias em geral.

4.2. Análise da confiabilidade do questionário

Com o intuito de verificar a confiabilidade do questionário proposto, foram considerados os resultados da aplicação deste na avaliação de uma das disciplinas do curso de Graduação em Engenharia de Produção da UENF, denominada "Computação para Engenharia II", segundo a percepção dos alunos. Os questionários foram distribuídos aos vinte e um alunos presentes ao final da aula, não estando o professor da disciplina presente durante a avaliação. Além disso, não foi permitida a troca de informações entre os alunos. Após um período de aproximadamente vinte minutos, os questionários foram recolhidos pelo aplicador.

Após a tabulação dos dados, foi possível determinar o coeficiente alfa, que apresentou valores altos em todas as Dimensões (vide tabela 2), ressaltando-se apenas o menor valor atribuído à Dimensão 6 (alfa = 0,743784). Entretanto, este valor não indica baixa confiabilidade do questionário à luz desta Dimensão.

TABELA 2
Confiabilidade do questionário à luz das Dimensões, segundo a percepção dos alunos

		ANÁLISE DE CONFIABILIDADE (COMPUTAÇÃO PARA ENGENHARIA II)			
Dimensão	Item	Alfa (α)	Corr. Inter-Item Médio	Alfa (se o item for eliminado)	Corr. Item-Total
D1	I ₁	0,872591	0,617479	0,900046	0,448148
	I ₂			0,802241	0,898365
	I ₃			0,865801	0,624897
	I ₄			0,828116	0,773939
	I ₅			0,818458	0,803425
D2	I ₆	0,891601	0,686162	0,891167	0,617092
	I ₇			0,904674	0,660975
	I ₈			0,858161	0,840809
	I ₉			0,867368	0,797411
	I ₁₀			0,850065	0,916376
	I ₁₁			0,871164	0,736563
D3	I ₁₂	0,944335	0,880768	0,872979	0,941134
	I ₁₃			0,979592	0,802897
	I ₁₄			0,897436	0,912774
D4	I ₁₅	0,857986	0,776513	0,927819	0,205371
	I ₁₆			0,826507	0,898996
	I ₁₇			0,836982	0,731324
	I ₁₈			0,832871	0,835079
	I ₁₉			0,850338	0,492996
	I ₂₀			0,845696	0,601337
	I ₂₁			0,830787	0,831175
	I ₂₂			0,832871	0,835079
	I ₂₃			0,830787	0,831175
	I ₂₄			0,834493	0,749663
D5	I ₂₅	0,857383	0,472102	0,840427	0,599102
	I ₂₆			0,845446	0,556250
	I ₂₇			0,825741	0,727488
	I ₂₈			0,851604	0,496525
	I ₂₉			0,822539	0,737220
	I ₃₀			0,848758	0,523428
	I ₃₁			0,829155	0,689387
	I ₃₂			0,852926	0,482234
D6	I ₃₃	0,743784	0,522037	0,703433	0,610441
	I ₃₄			0,691715	0,685341
	I ₃₅			0,711679	0,566083
	I ₃₆			0,700441	0,531805
	I ₃₇			0,659818	0,635139
	I ₃₈			0,810203	0,484261

Uma análise mais minuciosa pode ser feita considerando simultaneamente o valor da Correlação Item-Total (vide Apêndice II) e o valor de alfa para a Dimensão, caso um determinado item desta Dimensão

seja eliminado do questionário. Por exemplo, considerando a Dimensão 6, o item 38 apresenta o menor valor de correlação (0,484261) e, uma vez este seja eliminado, o alfa para esta Dimensão se elevará a 0,810203. Vale destacar que este item não necessariamente precisa ser imediatamente eliminado do questionário, pois a Dimensão 6 apresenta um valor de alfa significativamente alto. Entretanto, se observou que tal item não está altamente correlacionado com os demais itens que compõem a Dimensão. Uma possível causa para a baixa correlação deste item pode ser o fato deste estar mal redigido ou estar associado a dois ou mais conceitos.

Por outro lado, uma análise de mesma natureza permite identificar quais itens são considerados mais importantes por assegurar maior confiabilidade à Dimensão. Mais especificamente, tais itens não podem ser eliminados do questionário sob pena de redução da confiabilidade do questionário à luz desta Dimensão (nesta aplicação considerou-se, por exemplo, o Item 37).

Na observância dos resultados obtidos na avaliação desta disciplina, os valores de alfa foram considerados satisfatórios para todas as Dimensões estabelecidas, não sendo recomendada a alteração nos itens do questionário nesta oportunidade.

4.3. Apresentação dos resultados aos professores das disciplinas

Ao final da análise dos dados da avaliação das disciplinas curriculares, os resultados foram enviados aos respectivos professores responsáveis pelas disciplinas, tendo em vista o seu conhecimento, auto-avaliação e reflexão. Para tanto, foi desenvolvido um modelo de Relatório sintetizado que apresenta (vide Apêndice 1): a Média da avaliação da disciplina e a Média Geral de todas as disciplinas, ambas à luz de cada item (considerando uma escala numérica com valores de 2 a -2, associadas aos conceitos de desempenho de “Muito Bom” a “Muito Ruim” e, de “Concordo Totalmente” a “Discordo Totalmente”), um gráfico comparativo dos valores das Médias à luz de cada item, e a transcrição dos comentários e sugestões mais frequentes.

Em especial, os resultados apresentados neste relatório revelam que a disciplina avaliada “Computação para Engenharias II” apresenta desempenho superior ao desempenho médio das demais disciplinas à luz de quase todos os itens, com exceção dos itens I_{15} (pontualidade do professor) e I_{31} (tempo de estudo do aluno dedicado à disciplina). Além disso, destacam-se os resultados do Item 38, que se refere ao exercício (pelo aluno) de atividades em paralelo ao curso de engenharia. Segundo as médias relacionadas a este item, existe muita discordância dos alunos com este fato. De certa forma, tais resultados parecem coerentes, visto que o curso de Graduação em Engenharia de Produção da UENF é em regime integral de ensino (tarde/noite), fato este que restringe a atuação dos alunos em outras atividades.

5. Considerações sobre a implementação do modelo

Dentre as etapas que compuseram a estrutura do modelo proposto, a etapa de definição das Dimensões e os respectivos itens a estas pertencentes foi considerada a mais exaustiva e a que demandou maior tempo de execução. Neste contexto, buscou-se incorporar da forma mais abrangente quanto possível (no sentido de se adequar à avaliação de disciplinas das diferentes áreas do conhecimento) os aspectos mais relevantes e pertinentes à avaliação de disciplinas curriculares. Além disso, a redação dos itens foi

paulatinamente aperfeiçoada (após sugestões elaboradas pelos membros do Colegiado do curso de Graduação), com o intuito de constituir um instrumento de pesquisa (questionário) de fácil aplicação e simples para ser respondido.

Outro aspecto relevante consiste na forma de aplicação dos questionários. Em semestres anteriores, os questionários foram aplicados de duas diferentes formas:

- 1) Entregues aos alunos para serem preenchidos e posteriormente devolvimos.
- 2) Entregues por um aplicador aos alunos em sala de aula, com a presença do professor da disciplina avaliada.

Na primeira situação foi observado um percentual de devolução não tão satisfatório quanto o desejado (menos de 80%) e também uma elevada repetição de resultados de julgamentos do desempenho da maioria das disciplinas. Em particular, foi possível questionar se os resultados obtidos refletiam, realmente, a percepção de cada aluno ou se estes caracterizavam a busca de um consenso, provavelmente consolidado através da reunião de alunos com o objetivo de responder os questionários simultaneamente. Além disso, a repetição das sentenças apresentadas nos campos de questões abertas (comentários e sugestões) sugeriu a confirmação deste fato.

Por outro lado, na segunda situação foi observada uma melhora significativa do desempenho das disciplinas à luz da maior parte dos itens, além de uma redução, também significativa, da quantidade de respostas às questões abertas. Em relação a tais resultados, ficou assegurada a percepção individual de cada aluno. Por outro lado, foi possível constatar que a presença em sala de aula do professor da disciplina avaliada provavelmente inibiu o julgamento dos alunos.

Com o intuito de reduzir ou eliminar as deficiências identificadas nas duas situações supracitadas, no experimento descrito neste trabalho os questionários foram distribuídos a cada aluno da disciplina, sob a orientação de um aplicador e sem a presença do professor da disciplina em questão.

6. Restrições à implementação do modelo

Sob o ponto de vista do gerenciamento de uma IES, um procedimento de auto-avaliação institucional deve estar direcionado para a qualidade e para a eficiência do ensino. Em particular, o modelo de auto-avaliação proposto tem caráter pontual e abrangência micro, pois busca avaliar cada disciplina de determinado Curso à luz de vários critérios, segundo a percepção do corpo discente. Neste sentido, para que um modelo de avaliação desta natureza seja realmente eficaz, é preciso:

- 1) que os órgãos de chefia e de direção da Instituição de Ensino, o corpo docente e os alunos reconheçam a importância deste modelo, apóiem o seu desenvolvimento e a sua implementação, buscando avaliar conscientemente os resultados obtidos. Caso contrário, os formulários poderão não ser respondidos ou ser respondidos displicentemente (obtenção de informações incompletas ou inconsistentes), os professores avaliados poderão utilizar os resultados não satisfatórios para impor represálias nas turmas futuras, resultados negativos poderão ser utilizados pelas Coordenações dos Cursos com o intuito de penalizar professores, etc.;

- 2) que haja uma equipe/comissão designada para coordenar e implementar o modelo. Em especial, é recomendável que esta equipe seja composta por pessoas pertencentes a uma Gerência de Ensino e não sejam professores ou coordenadores de curso, não estando assim submetidas ao processo de avaliação em questão. Além de proporcionar uma melhor organização em todo o processo de avaliação, esta equipe pode evitar, dentre outros problemas, a manipulação fraudulenta dos resultados, e
- 3) que a identificação dos avaliadores seja optativa, reduzindo a influência de fatores externos. Em geral, avaliações deste tipo são mais realistas e sinceras.

Finalmente, é importante ressaltar que o modelo de auto-avaliação proposto pode ser empregue experimentalmente na avaliação das disciplinas de um único Curso de Graduação, sendo posteriormente estendido aos demais cursos. Neste período experimental, os procedimentos de aplicação devem ser detalhadamente registrados, permitindo o rastreamento da informação em caso de ocorrência de problemas.

7. Conclusões

Nos tempos atuais, a importância da Educação no desenvolvimento de um País é um fato inquestionável. Ao mesmo passo em que se estimula a disseminação do conhecimento pela sociedade, principalmente através da expansão do número de Instituições de Ensino Superior e do aperfeiçoamento dos docentes (cursos de tecnologia, mestrado e doutorado), é crucial assegurar a qualidade do ensino nestas Instituições.

Com o intuito de contribuir para a continuidade e aprimoramento das pesquisas focadas nos sistemas de avaliação do ensino superior, este trabalho propôs um modelo para a avaliação de disciplinas curriculares segundo a percepção dos alunos, à luz de vários critérios. Após a realização de um experimento, foi verificada a confiabilidade do instrumento de pesquisa (questionário) através do emprego do coeficiente alfa de Cronbach. Os valores de alfa obtidos foram satisfatórios, possibilitando a implementação do modelo proposto em outros experimentos.

Além disso, o relatório de avaliação que é encaminhado a cada professor responsável pela disciplina permite que tal professor reflita sobre o desempenho desta disciplina à luz das Dimensões e itens considerados. Em concordância com estes resultados, os professores poderão utilizar estas informações para aprimorar as disciplinas à luz dos itens considerados mais críticos.

Entretanto, é importante ressaltar que o modelo proposto não pretende substituir os instrumentos de avaliação desenvolvidos pelo MEC/INEP. Mais especificamente, este modelo se propõe atuar de forma complementar a tais instrumentos. Atuando no contexto da auto-avaliação de IES, o modelo proposto busca ressaltar a importância da avaliação "interna" da Qualidade do Ensino, onde as avaliações sob o ponto de vista dos alunos representam uma fonte de informações de grande relevância.

Bibliografia

CRONBACH, L. J. (1951): "Coefficient alpha and internal structure of tests", *Psychometrika*, 16, pp. 297-334.

- CRONBACH, L. J. (2004): "My current thoughts on coefficient alpha and successors procedures", *Educational and Psychological Measurement*, vol. 64, n.º 3, pp. 391-418.
- HAYES, B. E. (1995): *Measuring Customer Satisfaction – Development and Use of Questionnaires*, ASQC. INEP (2002): Manual de Avaliação Institucional – Centros Universitários. Versão preliminar.
- INEP (2006): Instituto Nacional de Estudos e Pesquisas Educacionais, www.inep.gov.br.
- MATTAR, F. N. (1996): *Pesquisa de Marketing*. São Paulo: Editora Atlas.
- OLIVEIRA, V. F.; BARBOSA, C. S. e CHRISPIM, E. M. (2005): Cursos de Engenharia de Produção no Brasil: Crescimento e Projeções. Anais do XXV ENEGEP. Porto Alegre, RS, Brasil.
- PETER, J. P. (1979): "Reliability: a review of psychometric basics and recent marketing practices", *Journal of Marketing Research*, vol. 16, pp. 6-17.
- SINAES (2003): Sistema Nacional de Avaliação da Educação Superior - *Bases Para Uma Nova Proposta de Avaliação Da Educação Superior*.
- TROCHIM, W. M. (2003): The Research Methods Knowledge Base, 3rd Edition. Internet WWW page, at URL: <<http://trochim.human.cornell.edu/kb/index.htm>>.

APÊNDICE 1: Modelo de questionário utilizado para a avaliação das disciplinas do curso de Graduação da Engenharia de Produção e Relatório com o resultado sintetizado da avaliação de cada disciplina



Formulário para a Avaliação das Disciplinas do Curso de Graduação em Engenharia de Produção (PERCEPÇÃO DOS ALUNOS)

Disciplina:	Professor:		
Aluno(a): (opcional)	Período atual:	CR acumulado:	
Ensino Médio:	<input type="checkbox"/> Escola Estadual	<input type="checkbox"/> Escola Federal	<input type="checkbox"/> Escola Particular

INDIQUE O GRAU DE DESEMPENHO DESTA DISCIPLINA EM RELAÇÃO ÀS SEGUINTE DIMENSÕES E ITENS:

Muito Bom	Bom	Neutro	Ruim	Muito Ruim
MB	B	N	R	MR

Dimensões e Itens

Grau de Desempenho				
MB	B	N	R	MR

Avaliação da Disciplina: conteúdo					
I ₁ Clareza na apresentação dos objetivos	MB	B	N	R	MR
I ₂ Cumprimento do programa proposto	MB	B	N	R	MR
I ₃ Carga horária para cumprir o programa proposto	MB	B	N	R	MR
I ₄ Existência exercícios resolvidos	MB	B	N	R	MR
I ₅ Existência de exercícios propostos (listas de exercícios)	MB	B	N	R	MR

Avaliação da Disciplina: provas e testes					
I ₆ Tempo de duração das provas	MB	B	N	R	MR
I ₇ Grau de dificuldade das questões das provas	MB	B	N	R	MR
I ₈ Adequação das questões das provas ao conteúdo da disciplina	MB	B	N	R	MR
I ₉ Adequação das questões das provas aos exercícios resolvidos e propostos	MB	B	N	R	MR
I ₁₀ Distribuição das questões quanto o conteúdo da disciplina	MB	B	N	R	MR
I ₁₁ Correção das questões	MB	B	N	R	MR

Avaliação do Material Didático					
I ₁₂ Divulgação e apresentação do material didático recomendado	MB	B	N	R	MR
I ₁₃ Disponibilidade do material didático recomendado	MB	B	N	R	MR
I ₁₄ Adequação do conteúdo do material didático ao conteúdo da disciplina	MB	B	N	R	MR

Avaliação do Professor da Disciplina					
I ₁₅ Pontualidade do professor	MB	B	N	R	MR
I ₁₆ Frequência do professor	MB	B	N	R	MR
I ₁₇ Organização no preparo das aulas	MB	B	N	R	MR
I ₁₈ Conhecimento do conteúdo da disciplina	MB	B	N	R	MR
I ₁₉ Clareza na apresentação e explicação do conteúdo	MB	B	N	R	MR
I ₂₀ Acessibilidade a esclarecimento de dúvidas dentro e fora da sala de aula	MB	B	N	R	MR
I ₂₁ Clareza na resolução de exercícios e esclarecimento de dúvidas	MB	B	N	R	MR
I ₂₂ Relacionamento professor-aluno	MB	B	N	R	MR
I ₂₃ Receptividade a divergências de opinião	MB	B	N	R	MR
I ₂₄ Receptividade a críticas e sugestões	MB	B	N	R	MR

Auto-avaliação (avaliar-se a si mesmo)					
I ₂₅ Pontualidade nas aulas	MB	B	N	R	MR
I ₂₆ Frequência nas aulas	MB	B	N	R	MR
I ₂₇ Participação nas aulas	MB	B	N	R	MR
I ₂₈ Interesse pela disciplina	MB	B	N	R	MR
I ₂₉ Acompanhamento do conteúdo das aulas	MB	B	N	R	MR
I ₃₀ Desempenho na resolução de exercícios (listas)	MB	B	N	R	MR
I ₃₁ Tempo de estudo na disciplina	MB	B	N	R	MR
I ₃₂ Desempenho nas provas e testes	MB	B	N	R	MR

Avaliação complementar	Discordo totalmente	Discordo	Neutro	Concordo	Concordo totalmente
I ₃₃ o conteúdo da disciplina foi adequado ao programa proposto					
I ₃₄ o conteúdo da disciplina foi apresentado adequadamente					
I ₃₅ as provas foram condizentes com o conteúdo apresentado					

I ₃₆	as notas obtidas refletem o esforço de estudo				
I ₃₇	o tempo de estudo na disciplina foi adequado				
I ₃₈	exerço alguma atividade em paralelo ao curso de engenharia				

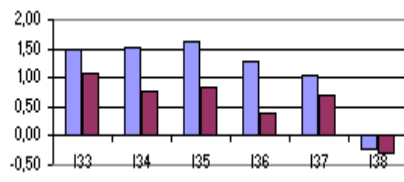
Utilize o espaço abaixo e o verso da folha para proferir comentários adicionais:

Comentários sobre conteúdo da disciplina e material didático:

Comentários sobre o professor da disciplina:

Sugestões:

		Resultado da Avaliação das Disciplinas do Curso de Graduação em Eng. de Produção (Percepção dos Alunos)				
Disciplina: Computação para Engenharia II		Prof.:				
Escala para avaliação do Grau de Desempenho da disciplina relação às Dimensões e critérios:						
Grau de Desempenho		Muito Bom (2)	Bom (1)	Neutro (0)	Ruim (-1)	Muito Ruim (-2)
Dimensões e Itens			Média	Média*	Grau de Desempenho Comparativo	
D1: Avaliação da Disciplina: conteúdo						
I ₁	Clareza na apresentação dos objetivos		1,67	0,84		
I ₂	Cumprimento do programa proposto		1,71	1,20		
I ₃	Carga horária para cumprir o programa proposto		1,71	1,18		
I ₄	Existência exercícios resolvidos		1,62	0,41		
I ₅	Existência de exercícios propostos (listas de exercícios)		1,67	0,67		
D2: Avaliação da Disciplina: provas e testes						
I ₆	Tempo de duração das provas		1,52	0,86		
I ₇	Grau de dificuldade das questões das provas		1,38	0,80		
I ₈	Adequação das questões das provas ao conteúdo da disciplina		1,67	0,95		
I ₉	Adequação das questões das provas aos exercícios resolvidos e propostos		1,76	0,75		
I ₁₀	Distribuição das questões quanto o conteúdo da disciplina		1,71	0,94		
I ₁₁	Correção das questões		1,62	0,74		
D3: Avaliação do Material Didático						
I ₁₂	Divulgação e apresentação do material didático recomendado		1,14	0,95		
I ₁₃	Disponibilidade do material didático recomendado		1,00	0,60		
I ₁₄	Adequação do conteúdo do material didático ao conteúdo da disciplina		1,19	1,00		
D4: Avaliação do Professor da Disciplina						
I ₁₅	Pontualidade do professor		0,81	0,92		
I ₁₆	Freqüência do professor		1,76	1,10		
I ₁₇	Organização no preparo das aulas		1,76	0,99		
I ₁₈	Conhecimento do conteúdo da disciplina		1,81	1,28		
I ₁₉	Clareza na apresentação e explicação do conteúdo		1,71	0,68		
I ₂₀	Acessibilidade a esclarecimento de dúvidas dentro e fora da sala de aula		1,48	0,86		
I ₂₁	Clareza na resolução de exercícios e esclarecimento de dúvidas		1,76	0,76		
I ₂₂	Relacionamento professor-aluno		1,81	1,08		
I ₂₃	Receptividade a divergências de opinião		1,76	0,98		
I ₂₄	Receptividade a críticas e sugestões		1,71	0,97		
D5: Auto-avaliação (avaliar-se a si mesmo)						
I ₂₅	Pontualidade nas aulas		1,10	1,05		
I ₂₆	Freqüência nas aulas		1,24	1,00		
I ₂₇	Participação nas aulas		1,10	0,77		
I ₂₈	Interesse pela disciplina		1,33	0,83		
I ₂₉	Acompanhamento do conteúdo das aulas		1,00	0,82		
I ₃₀	Desempenho na resolução de exercícios (listas)		1,05	0,86		
I ₃₁	Tempo de estudo na disciplina		0,48	0,62		
I ₃₂	Desempenho nas provas e testes		1,10	0,60		

Grau de Concordância	Discordo totalmente(-2)	Discordo(-1)	Neutro (0)	Concordo (1)	Concordo totalmente(2)
D6: Avaliação complementar					
I ₃₃	o conteúdo da disciplina foi adequado ao programa proposto		1,48	1,06	
I ₃₄	o conteúdo da disciplina foi apresentado adequadamente		1,52	0,77	
I ₃₅	as provas foram condizentes com o conteúdo apresentado		1,62	0,83	
I ₃₆	as notas obtidas refletem o esforço de estudo		1,29	0,40	
I ₃₇	o tempo de estudo na disciplina foi adequado		1,05	0,71	
I ₃₈	exerço alguma atividade em paralelo ao curso de engenharia		-0,24	-0,29	
Comentários sobre conteúdo da disciplina e material didático:			Não há material didático disponível;		
Faltam computadores para as aulas práticas;					
Comentários sobre o professor da disciplina:			O professor é excelente; Poderia ter trabalhado mais conteúdo no período;		
Sugestões:					
Aumentar o número de computadores para o melhor desempenho do aluno e para que todos possam praticar o que o professor ensina;					

APÊNDICE 2

COEFICIENTE DE CORRELAÇÃO INTER-ITEM: este coeficiente de correlação permite determinar o grau em que cada item do questionário está relacionado com os demais itens que avaliam a Dimensão a que este item pertence. Sejam X e Y dois itens que avaliam certa Dimensão. O coeficiente de correlação Inter-Item é dado por:

$$\rho = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^n (x_i - \mu_Y)(y_i - \mu_X)}{n \sigma_X \sigma_Y}, \quad (A.1)$$

onde:

- n é o número de avaliadores;
- μ_X é a média dos valores do Item X ;
- μ_Y é a média dos valores do Item Y ;
- σ_X é o desvio padrão dos valores do Item X ;
- σ_Y é o desvio padrão dos valores do Item Y .

COEFICIENTE DE CORRELAÇÃO ITEM-TOTAL: este coeficiente de correlação permite determinar o grau em que cada item do questionário está relacionado à soma total dos pontos (pontuação total) obtidos pelos demais itens que avaliam a Dimensão a que este item pertence (é importante ressaltar que a pontuação do item correlacionado não deve ser considerada na pontuação total).