

MODELO MULTICRITÉRIO NA AVALIAÇÃO DOCENTE

Nelson Hein

Universidade Regional de Blumenau – FURB
Rua Antônio da Veiga, 140 – Blumenau-SC
hein@furb.br

Adriana Kroenke

Universidade Federal do Paraná – UFPR
Universidade Regional de Blumenau – FURB
Rua Antônio da Veiga, 140 – Blumenau-SC
didlen@terra.com.br

RESUMO

O artigo que se apresenta é resultado do projeto piloto na avaliação docente da Universidade Regional de Blumenau. Nele é discutido o projeto elaborado pela equipe de avaliação interna da IES, discutindo as dimensões, questões e técnicas quantitativas de análise e formação do Índice FURB de Avaliação Docente (IFAD). As ferramentas multicritério utilizadas foram a teoria do consenso, entropia da informação e o método de análise decisória multicritério TOPSIS. Foram avaliados dois cursos de graduação, engenharia química e pedagogia, totalizando 56 professores avaliados por um conjunto de 1979 alunos. O índice IFAD mínimo dentre os docentes do curso de Engenharia Química está em 1,53 (PR37) e o máximo em 4,48 (PR9). No curso de Pedagogia o mínimo avaliado foi o IFAD do PR58 com 1,43 e a maior nota avaliada foi a do PR49 com IFAD=4,54.

Palavra-chave: Avaliação docente. Ferramentas multicritério. TOPSIS.

ABSTRACT

The present article is the pilot project result on teacher evaluation of the Regional University of Blumenau (FURB). In it is discussed the draft prepared by the team of internal evaluation of educational institution, discussing the dimensions, issues and quantitative analysis techniques and development of FURB's Teacher Evaluation Index (IFAD). The multicriteria tools used were the consensus theory, information entropy and method of multiple criteria decision analysis TOPSIS. Were assessed two undergraduate courses, Chemical Engineering and Pedagogy, totaling 56 teachers evaluated by a group of 1979 students. The IFAD minimum rate among Chemical Engineering teachers is 1.53 (PR37) and the maximum 4.48 (PR9). In Pedagogy course, the minimum rated was the IFAD of PR58 resulting 1.43 and highest grade assessed was PR49 with IFAD = 4.54.

Key-words: Teacher evaluation. Multicriteria tools. TOPSIS.

1. INTRODUÇÃO

A responsabilidade das universidades, frente aos desafios e metas exige uma auto avaliação do cotidiano do trabalho, da docência, da gestão de pesquisa e extensão que nela se realizam. Esta auto avaliação pressupõe a participação de toda comunidade acadêmica, que formula juízos sobre o valor da qualidade institucional.

Ente importante na qualidade de uma universidade é o seu corpo docente. A avaliação interna desse corpo espelha a perspectiva institucional futura, na qual a ação docente é um dos seus elementos. Como destaca o Programa de Avaliação das Universidades Brasileiras (PAIUD), essa avaliação não deve ser um processo punitivo, nem mesmo de coibição, mas deve apresentar indicadores com vistas à qualidade (BRASIL, 1994; BRASIL, 2004).

Este projeto apresenta os resultados de um estudo que diagnostica e avalia o gerenciamento de ensino a partir da visão dos alunos de Engenharia de Química e Pedagogia da Universidade Regional de Blumenau (FURB), localizada no Vale do Itajaí, em Blumenau (SC). O projeto e seus resultados são fruto de um programa piloto, que hoje é institucionalizado plenamente aos demais cursos da FURB, para um total de 40 cursos, 842 professores e cerca de 10200 alunos.

Para a avaliação, foram sanadas uma série de problemas que eram questionados anteriormente pelos professores avaliados: significância da amostra e coerência das respostas assinaladas pelos alunos (avaliadores). Para a significância estatística de dados foi calculado o tamanho da amostra para cada grupo de alunos, divididos em intervalos, para cada disciplina avaliada (podendo um docente ser avaliados por mais de um grupo). Para resolver o problema da coerência foi introduzido o conceito de consistência da média por questão gerando a informação à avaliação e da entropia, para gerar o peso da informação.

Tudo isso resulta em um indicador, denominado Índice FURB de Avaliação Docente (IFAD), limitado em uma escala entre 1 a 5, similar a adotada pelo INEP (Instituto Nacional de Ensino e Pesquisa). Para a obtenção do referido índice foram utilizadas técnicas de análise decisória multicritério, além de outras técnicas quantitativas.

2. REFERENCIAL TEÓRICO

Os primeiros projetos de avaliação das IES desenvolvidos pelas universidades brasileiras surgiram entre as décadas de 80 e 90, a partir da necessidade de se avaliar a universidade, devido a críticas divulgadas nos meios de comunicação do país em relação a essas instituições (BOTH, 2005). A partir desse momento, foi consagrado o termo “avaliação institucional”. Desse modo, o movimento avaliativo se desenvolveu pela necessidade de se provocar melhorias na gestão universitária.

A questão da avaliação institucional levou ao surgimento, em 1993, do Programa de Avaliação Institucional das Universidades Brasileiras (PAIUB), que é o instrumento utilizado pelo governo para definir um projeto de financiamento de projetos de avaliação institucional. Nesse instrumento são estabelecidos os princípios e os objetivos da avaliação institucional brasileira, de forma que cada universidade possa desempenhar as atividades na busca da melhoria da qualidade (BRASIL, 1994).

Em 1994, o PAIUB estabeleceu sete princípios norteadores do processo de avaliação: globalidade, comparabilidade, respeito à identidade institucional, não premiação ou punição, adesão voluntária, legitimidade do processo e continuidade do processo. Percebe-se, então, que devem ser observadas as características próprias de cada instituição, consideradas as forças atuantes em seu ambiente específico, a missão institucional e seu projeto institucional. Entendendo-se que a avaliação institucional caracteriza-se como um processo sistêmico e contínuo que possibilita a reflexão sobre as atividades desenvolvidas, permitindo o aperfeiçoamento permanente da qualidade do ensino superior.

Posteriormente ao PAIUB, o Decreto nº 2.036 de 10 de outubro de 1996 estabeleceu procedimentos para a avaliação dos cursos das instituições de ensino superior.

Depois disso, em 14 de abril de 2004, foi criado, por meio da Lei nº 10.861, o Sistema Nacional de Avaliação da Educação Superior (BRASIL, 2004) que se fundamenta na necessidade de promover a melhoria da qualidade de educação superior. Além de fornecer orientação e expansão da sua oferta, o aumento permanente da sua efetividade acadêmica social e, especialmente, do aprofundamento dos seus compromissos e responsabilidades sociais.

De acordo com o documento do INEP/SINAES (BRASIL, 2004) a avaliação das instituições de educação superior tem caráter formativo, pois visa ao aperfeiçoamento dos agentes da comunidade acadêmica e da instituição como um todo.

A auto avaliação na FURB iniciou-se em á partir de 1995. Desde então a instituição compromete-se com o processo de auto avaliação, não apenas por um momento, mas sim de caráter constante e renovável, tal fato permite uma praticamente global da instituição o facilita o redimensionamento da política e dos planos para a universidade.

A avaliação dos professores iniciou na mesma época, sendo o modelo modificado de tempos em tempos, porém sem nenhum grande avanço em sua análise.

3. MATERIAIS E MÉTODOS

A pesquisa é caracterizada como exploratória, com procedimento bibliográfico, documental e de levantamento de dados. A análise e interpretação dos dados são de natureza quantitativa. O projeto piloto envolveu uma população constituída pelos 1979 alunos e 56 professores da Universidade Regional de Blumenau. A avaliação foi realizada tomando 58 turmas, das disciplinas oferecidas no 1º semestre de 2013, dos cursos de Engenharia Química e Pedagogia.

O questionário é composto por duas dimensões: (a) **social e ética** e (b) **ensino e aprendizagem**, distribuídas em 19 perguntas, assim organizadas:

(a) Dimensão Social e Ética

Atitude docente. Ao longo do período o (a) professor (a)?

- 1) Estabeleceu uma relação de cortesia, respeito e atenção com os estudantes.
- 2) Incentivou os estudantes ao questionamento dos fundamentos, teorias, conceitos, etc.
- 3) Estimulou os estudantes a estabelecerem conclusões e participar de forma efetiva nas discussões, exercícios e atividades.
- 4) Estimulou os estudantes a integrarem conhecimentos com outras disciplinas correlatas.
- 5) Compareceu às aulas previstas, cumprindo o horário estabelecido para a disciplina.
- 6) Registrou a pontualidade e frequência dos alunos.
- 7) Estimulou os estudantes a valorizarem o curso, divulgando uma boa imagem da instituição.
- 8) Demonstrou estar atualizado com as tendências da área da disciplina citando autores, eventos, experiências, outros profissionais.

(b) Dimensão ensino e Aprendizagem

Quanto a organização das ações de ensino-aprendizagem o (a) professor (a)?

- 9) Apresentou e discutiu o Plano de Ensino-Aprendizagem da disciplina.
- 10) Deixou claro o(s) objetivo(s) da disciplina e das unidades de ensino.
- 11) Cumpriu o conteúdo da disciplina de acordo com Plano de Ensino-Aprendizagem da disciplina.

- 12) Demonstrou clareza, preparação e objetividade na explicação do conteúdo da disciplina.
- 13) Apontou a relevância e/ou aplicação do conteúdo estudado.
- 14) Indicou fontes de consulta (literatura, sites, soft ...) adequadas à proposta da disciplina.
- 15) Exigiu nas avaliações de aprendizagem os conteúdos desenvolvidos de acordo com os objetivos do Plano de Ensino-Aprendizagem.

Na realização das ações de ensino-aprendizagem o (a) professor (a)?

- 16) Elaborou e executou instrumentos e procedimentos de avaliação conforme previsto no Plano de Ensino-Aprendizagem.
- 17) Utilizou procedimentos didáticos diversificados e adequados ao(s) objetivo(s) da disciplina.
- 18) Utilizou instrumentos de avaliação diversificados e adequados ao(s) objetivo(s) da disciplina.
- 19) Analisou e comentou com os alunos os resultados de provas, testes e outras atividades desenvolvidas, corrigindo erros e esclarecendo dúvidas.

Foram analisadas somente as turmas que alcançam amostras com 95% de confiança. O esquema de para dar legitimidade será dado na forma:

Tabela 01 – Tamanho da amostra por classe

# Alunos	# Amostra
$2 < N \leq 10$	N-1
$10 < N \leq 20$	N-2
$20 < N \leq 30$	N-3
$30 < N \leq 40$	N-4
$40 < N \leq 50$	N-6
$50 < N \leq 60$	N-8
$60 < N \leq 70$	N-10
$70 < N \leq 80$	N-13

Fonte: Elaborado pelos autores.

A escala usada é ordinal de Likert de 5 pontos, assim pontuando: Discordo Totalmente (1), Discordo (2), Nem concordo, nem discordo (3), Concordo (4) e Concordo Totalmente (5). Mesmo a escala apresentando uma série de limitações técnicas, não se pode lhe tirar o mérito em ser conhecida pelos alunos e por apresentar um ponto neutro, no caso da escala de 5 pontos o valor '3', que pode ser entendido como sendo 0 (zero) em uma escala de -2 a +2.

Além do uso da escala foi incorporada a figura do consenso avaliado pela expressão:

$$CONS(X) = 1 + \sum_{i=1}^n p_i \log_2 \left(1 - \frac{|X_i - \mu_X|}{d_X} \right)$$

Onde X é uma variável ordinal (questão); p_i é o percentual associado a cada X_i ; d_X é a amplitude da escala, que neste caso $d_X = 4$ e μ_X é a média. O consenso dará respaldo a média obtida, pois servirá de critério de variabilidade dela. Segundo Wierman e Tastle (2007) o consenso deve ser assim interpretado.

Tabela 02 – Classificação do grau de consenso

Intervalo	Classificação
$\text{CONS}(X) \geq 90\%$	Consenso Muito Forte
$80\% \leq \text{CONS}(X) < 90\%$	Consenso Forte
$60\% \leq \text{CONS}(X) < 80\%$	Consenso Moderado
$40\% \leq \text{CONS}(X) < 60\%$	Equilíbrio
$20\% \leq \text{CONS}(X) < 40\%$	Dissenso Moderado
$10\% \leq \text{CONS}(X) < 20\%$	Dissenso Forte
$\text{CONS}(X) < 10\%$	Dissenso Muito Forte

Fonte: Wierman e Tastle (2007).

Os dados na simulação do modelo a seguir foram coletados na pesquisa realizada junto ao corpo discente. Contudo, são parte da investigação e servem unicamente para oferecer uma compreensão algorítmica e numérica do projeto de avaliação docente.

A seguir é apresentada uma simulação em que são avaliados 10 professores, sendo cada um deles foi avaliado por 10 alunos, em um questionário de 8 perguntas. Para exemplificar, apresenta-se um quadro com as respostas simuladas para um entendimento da técnica.

Quadro 01 - Relatório das respostas às questões dados pelos alunos

Aluno	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
A1	3	2	1	1	1	5	4	1
A2	4	3	4	2	1	4	4	1
A3	5	4	5	3	2	3	5	1
A4	4	3	4	4	1	4	5	2
A5	5	2	5	5	1	3	5	1
A6	4	3	4	4	5	2	5	1
A7	5	4	5	3	3	2	5	1
A8	3	2	4	2	5	3	5	1
A9	2	1	3	1	5	4	5	1
A10	3	2	4	2	5	3	5	1
Média	3,8	2,6	3,9	2,7	2,9	3,3	4,9	1,1
Cons(X)	0,64	0,66	0,62	0,50	0,16	0,68	0,93	0,93

Fonte: Dados parciais da pesquisa.

As colunas das perguntas Q7 e Q8, possuem ambas consenso de 93%. Contudo, deve-se entender que o “consenso muito forte” impacta sobre a média, pois ratifica a alta pontuação recebida na pergunta Q7 e a baixa nota recebida em Q8. Observe-se que na pergunta Q4, com média 2,7 o consenso foi de apenas 50% e na pergunta Q5 a média em 2,9 vem acompanhada de um índice de consenso de apenas 16%. A pontuação total do professor em cada questão será dada pelo produto da média pelo seu consenso: Valor = Média \times Consenso. No caso deste professor sua pontuação ficará assim determinada.

Quadro 02: Valor associado a cada questão ao professor em avaliação

Prof	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Valor	2,43	1,72	2,42	1,34	0,46	2,24	4,56	1,02

Fonte: Dados da pesquisa

Supondo agora que o quadro a seguir apresente os valores obtidos por consenso dos 10 professores avaliados. Além desta análise cada questão recebeu um fator de importância, calculado pela entropia (ZELENY, 1982) presente em cada lote de respostas, ou seja, foram usadas ferramentas da teoria da informação, calculada por:

$$ENT(X) = -\alpha \sum_{k=1}^n \frac{x_i^k}{X_i} \ln \left(\frac{x_i^k}{X_i} \right)$$

Quadro 03 - Valor das questões dos professores avaliados

Profes	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
PR1	2,43	1,72	2,42	1,34	0,16	2,24	4,56	1,02
PR2	3,46	2,33	3,42	3,45	1,23	4,78	4,32	2,34
PR3	3,78	2,67	3,78	4,02	1,56	4,56	3,46	4,31
P44	2,45	3,42	2,25	1,24	0,67	4,02	3,21	3,67
PR5	1,22	1,78	1,23	2,34	1,54	3,45	3,56	3,21
PR6	4,02	3,45	4,56	3,21	1,56	2,32	3,56	2,89
PR7	3,57	3,42	3,33	2,67	0,98	1,21	3,01	2,56
PR8	2,41	3,56	2,67	3,54	1,89	0,56	2,23	3,12
PR9	4,12	2,42	3,45	3,21	2,03	0,34	0,89	3,42
PR10	3,51	4,67	2,67	1,78	2,12	1,45	3,76	2,65
Max	4,12	4,67	4,56	4,02	2,12	4,78	4,56	4,31
Min	1,22	1,72	1,23	1,24	0,16	0,34	0,89	1,02
ENT(X)	0,980	0,981	0,980	0,972	0,950	0,907	0,975	0,980
Inf%	7,1	6,9	7,4	10,0	18,3	33,7	9,0	7,4

Fonte: Dados da Pesquisa.

O valor da informação, ou seja, o peso de cada questão é dado pela parcela da entropia no total das 8 questões: $\omega_i = \frac{1-e(P_i)}{n-E}$, $e(P_i)$ é a entropia de cada questão, $n = 8$ pois é o total da entropia possível e $E = 7,725$, que foi a entropia gerada pelas 8 questões.

Gerada a matriz das pontuações dos professores, o problema torna-se um problema de análise multicritério, onde se pretende ranquear o grupo de professores, usando como critérios o desempenho em cada uma das questões que foram respondidas.

O modelo de análise multicritério a ser utilizado é o TOPSIS, acrônimo de *Technique for Order Preference by Similarity to Ideal Solution*. O modelo inicia com uma matriz de decisão composta por alternativas (professores) e critérios (questões).

$$A = \begin{bmatrix} v_{11} & \dots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{m1} & \dots & v_{mn} \end{bmatrix}$$

A técnica possui os seguintes passos:

1º Passo: Cálculo das soluções ideais positivas A^+ (benefícios) e das soluções ideais negativas A^- (custos), da seguinte forma:

$A^+ = (p_1^+, p_2^+, \dots, p_m^+)$ e $A^- = (p_1^-, p_2^-, \dots, p_m^-)$,
onde:

$$p_j^+ = \{Max_i p_{ij}, j \in J_1; Min_i p_{ij}, j \in J_2\}$$

$$p_j^- = \{Min_i p_{ij}, j \in J_1; Max_i p_{ij}, j \in J_2\},$$

onde: J_1 e J_2 representam respectivamente o critério benefício e custo. Na questão em estudo o objetivo é a busca de benefícios, ou seja, quanto mais alto o desempenho na escala Likert melhor.

2º Passo: O cálculo das distâncias euclidianas entre os benefícios é feito na forma:

$d^+ = \sqrt{\sum_{j=1}^n w_j (p_j^+ - p_{ij})^2}$, com $i=1, \dots, m$ e $d^- = \sqrt{\sum_{j=1}^n w_j (p_j^- - p_{ij})^2}$, com $i=1, \dots, m$. Os valores de w_i , nada mais são que o grau de importância de cada questão, obtido pela entropia de cada uma delas.

3º Passo: Cálculo da proximidade relativa $\xi_i = \frac{d_i^-}{d_i^+ + d_i^-}$.

Exemplificando a situação com o questionário dos 10 professores, com os valores obtidos pela média e consenso.

Quadro 04 - Resultado Final da Simulação e atribuição do IFAD.

Prof	V1	V2	V3	V4	V5	V6	V7	V8	d_i^+	d_i^-	ξ_i	IFAD
PR1	2,43	1,72	2,42	1,34	0,16	2,24	4,56	1,02	2,35	4,09	0,634	3,54
PR2	3,46	2,33	3,42	3,45	1,23	4,78	4,32	2,34	0,99	6,15	0,861	4,45
PR3	3,78	2,67	3,78	4,02	1,56	4,56	3,46	4,31	0,72	7,10	0,908	4,63
PR4	2,45	3,42	2,25	1,24	0,67	4,02	3,21	3,67	1,50	4,97	0,768	4,07
PR5	1,22	1,78	1,23	2,34	1,54	3,45	3,56	3,21	1,77	4,04	0,696	3,78
PR6	4,02	3,45	4,56	3,21	1,56	2,32	3,56	2,89	1,58	6,69	0,808	4,23
PR7	3,57	3,42	3,33	2,67	0,98	1,21	3,01	2,56	2,32	5,14	0,689	3,75
PR8	2,41	3,56	2,67	3,54	1,89	0,56	2,23	3,12	2,68	4,82	0,643	3,57
PR9	4,12	2,42	3,45	3,21	2,03	0,34	0,89	3,42	2,90	5,33	0,647	3,59
PR10	3,51	4,67	2,67	1,78	2,12	1,45	3,76	2,65	2,19	5,95	0,731	3,92

Fonte: Dados da pesquisa.

Acrescente-se que o modelo gera ao final uma nota na escala adotada pelo INEP para atribuição de Conceito Preliminar de Curso (CPC) e o Índice Geral de Cursos (IGC) é dada pelo intervalo contínuo [1, 5], que no projeto piloto recebeu o nome de IFAD (Índice FURB de Avaliação Docente).

4. ANÁLISE DOS RESULTADOS

Os questionários foram respondidos de forma online, formando uma base de dados em Access, extraídos estes dados são gerados relatórios, após passar por todas as técnicas já apresentadas.

O IFAD de cada um dos 58 professores, dos dois cursos em análise foram os que aparecem nos Quadros 5 e 6.

Quadro 5: IFAD dos docentes do curso de Engenharia Química

Prof.	IFAD								
PR1	2,61	PR11	2,80	PR21	3,51	PR31	3,17	PR41	3,40
PR2	3,28	PR12	2,98	PR22	3,03	PR32	2,89	PR42	3,66
PR3	3,20	PR13	3,91	PR23	3,15	PR33	3,00	-	-
PR4	3,27	PR14	2,57	PR24	4,33	PR34	3,80	-	-
PR5	3,09	PR15	1,82	PR25	2,90	PR35	1,59	-	-
PR6	4,42	PR16	3,33	PR26	3,61	PR36	1,91	-	-
PR7	2,94	PR17	2,65	PR27	3,92	PR37	1,53	-	-
PR8	2,80	PR18	1,97	PR28	3,03	PR38	3,80	-	-
PR9	4,48	PR19	3,97	PR29	2,13	PR39	3,02	-	-
PR10	2,45	PR20	2,75	PR30	3,62	PR40	2,86	-	-

Fonte: Dados da pesquisa.

Quadro 6: IFAD dos docentes do curso de Pedagogia

Prof.	IFAD	Prof.	IFAD	Prof.	IFAD	Prof.	IFAD
PR43	4,04	PR47	3,44	PR51	2,74	PR55	2,22
PR44	1,96	PR48	4,34	PR52	3,22	PR56	3,34
PR45	4,42	PR49	4,54	PR53	1,51	PR67	3,40
PR46	4,13	PR50	4,09	PR54	3,98	PR58	1,43

Fonte: Dados da Pesquisa

O índice IFAD mínimo dentre os docentes do curso de Engenharia Química está em 1,53 (PR37) e o máximo em 4,48 (PR9). No curso de Pedagogia o mínimo avaliado foi o IFAD do PR58 com 1,43 e a maior nota avaliada foi a do PR49 com IFAD = 4,54.

5. CONCLUSÕES

O modelo que está em uso na Universidade Regional de Blumenau é fruto de muitas tentativas anteriores em melhorar o sistema de auto avaliação. O modelo apresentado surgiu do questionamento dos professores que semestralmente são avaliados e a divulgação dos resultados causava constrangimento e descontentamento por parte dos envolvidos.

O uso apenas da média aritmética impossibilitava uma melhor análise. Com o uso dos conceitos de consenso, peso da informação e especificamente a introdução do IFAD reduziu queixas e serve atualmente para a escalção dos professores aos cursos da universidade, servindo de ferramenta de auxílio à decisão por parte de coordenadores e chefes de departamentos.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] BOTH, Ivo. Da avaliação precursora ao comportamento da avaliação: PAIUB e SINAES construindo pontes. *Revista da Rede de Avaliação Institucional da Educação Superior*, n. 4, p. 61-73, 2005.
- [2] BRASIL. *Programa de avaliação das universidades brasileiras (PAIUB)*. Brasília: Ministério da Educação, Secretaria de Educação Superior, 1994.
- [3] BRASIL. *Sistema nacional de avaliação da educação superior (SINAES)*. Brasília: Ministério da Educação (MEC), Secretaria de Educação Superior, 2004.
- [4] WIERMAN, M. J.; TASTLE, W. J.. Consensus and dissention – a measure of ordinal dispersion. *International Journal of Approximate Reasoning*. New York, pp. 531 – 545, 2007.

- [5] ZELENY, Milan. *Multiple Criteria Decision Making*. New York:Mc Graw-Hill, 1982.