

Peer evaluation: Assegurando a qualidade na avaliação em escala

Ribeirão Preto - SP - Abril 2013

Ronai Ferreira Ramos - UNIVESP/USP - ronai@usp.br

José Dutra de Oliveira Neto - FEA-RP/USP - dutra@usp.br

David Forli inocente- INEPAD - forli@inepad.org.br

Pesquisa e Avaliação Educação Superior Macro D / Meso L / Micro N Descrição de Projeto em Andamento Investigação Científica RESUMO

*Atualmente no Brasil a educação a distância é uma modalidade de ensino em constante crescimento. Algumas alternativas viáveis economicamente como peer evaluation permitem a avaliação de questões dissertativas com um custo relativamente baixo e sua eficácia é suportada por diversos estudos. O objetivo desta pesquisa é avaliar a adoção da metodologia de autoavaliação e avaliação por pares, comparando-as com a avaliação tradicional realizada por tutores, quando as turmas possuem um elevado número de alunos. Para a realização deste estudo pretende-se utilizar o Workshop do Moodle aplicado em uma turma de 200 alunos de um curso de especialização na modalidade virtual. A modelagem do trabalho seguirá praticamente quatro etapas: treinamento dos grupos; disponibilização e execução das atividades no ambiente virtual, com prazos para início e término; Correção das atividades e atribuição de notas, garantindo que os alunos possam se autoavaliar e os pares acessem as atividades dos colegas (aleatoriamente no estilo *blind review*). Por fim uma avaliação geral dos dados e, possível, discrepância entre os três instrumentos de avaliação comparados (tutores, autoavaliação e pares) será realizada e testada por meio de análises estatísticas.*

Palavras chave: autoavaliação, feedback, *peer evaluation*

1 Introdução

Atualmente no Brasil a educação a distância é uma modalidade de ensino em constante crescimento, não só para cursos de graduação, mas também na pós-graduação (*lato sensu*)[1]. Este aumento exponencial de ofertas, leva a uma preocupação com a qualidade dos cursos.

Na busca pela melhoria contínua da qualidade no ensino, surge a busca por novos métodos de ensino e estratégias educacionais. Uma das formas de mensurar esta qualidade é por meio do processo de avaliação. No Brasil, independentemente do caráter educacional ser conteudista ou construtivista, é muito comum identificarmos no processo de avaliação para as atividades online, desenvolvidas pelos discentes, a correção centralizada pela equipe de tutores.

Em função do modelo econômico e financeiro geralmente adotado nas instituições de ensino e o grande número de alunos, o custo de uma avaliação torna-se um elemento com grande impacto no resultado econômico do curso. Diante deste fato, surge as alternativas de avaliação usando testes de múltipla escolha, que embora possa trazer grandes benefícios econômicos para cursos em grande escala, não pode deixar de evitar o viés que pode ocorrer, já que a resposta não é espontânea [2].

Outra forma de avaliação seria as questões dissertativas que permitem ao aluno expressar sua resposta sem influência do instrutor, além de permitir o desenvolvimento do pensamento mais crítico do aluno. Esta opção, porém tem dois problemas: custo que varia em função do número de alunos e tempo de resposta (*feedback*) muito grande.

Os custos dos tutores representam uma grande parcela das despesas de cursos online e podem impactar negativamente nos resultados econômicos. Já o *feedback* deve ser rápido para tornar-se útil e significativo de tal maneira que os alunos possam fazer mudanças em suas próxima tentativa [3]. Este mesmo autor destaca que o valor do *feedback* decresce a medida que aumenta o tempo entre a realização da atividade e o recebimento do *feedback*.

Diante deste cenário, surge uma alternativa a estratégia do *peer evaluation* que permite a avaliação de questões dissertativas de maneira rápida e com um custo relativamente baixo.

O *peer evaluation* é uma estratégia efetiva no aprendizado colaborativo que permite os alunos avaliarem os seus colegas em uma determinada atividade. Diversos trabalhos dão credibilidade a estratégia educacional *peer evaluation* [4][5].

Este tipo de estratégia que facilita a interação entre os alunos pode maximizar a participação do aluno colaborando para a melhoria de suas competências sociais que são desenvolvidas durante o processo de construção do conhecimento acadêmico [4]. O mesmo autor acredita que o seu pouco uso é devido a sua credibilidade, já que se acredita que as avaliações dos alunos são consideradas menos confiáveis que a dos tutores. Este fato decorre das diferentes visões entre alunos e tutores que podem ser melhor alinhadas ao utilizarmos as rubricas de avaliação. As rubricas permitem que ambos os atores utilizem o mesmo critério para responder e avaliar a atividade[6].

Uma das críticas ao processo desta estratégia é o fato de muitos dos estudos não utilizarem do formato *blind review* o que pode resultar notas inflacionadas ou deflacionadas no resultado final.

Embora existam algumas abordagens possíveis para realizar estudos e caracterizar, de forma comparativa, se há diferença significativa na atribuição de notas entre alunos e tutores, pouco se encontrou na literatura sobre a utilização desta prática no Brasil. Contudo, um trabalho [7] destaca problemas na avaliação colaborativa (entre pares) pela falta, em alguns casos, de um *feedback* adequado. Entretanto, o estudo aqui proposto tem como principal objetivo avaliar, a adoção da metodologia de autoavaliação e avaliação por pares, comparando-as com a avaliação tradicional realizada por tutores, e destacando-as como formas viáveis para o desenvolvimento pedagógico e o melhor aproveitamento das equipes de tutores. O presente estudo utiliza-se do *blind review* e da rubrica no processo avaliativo.

Deste modo, com a realização do estudo pretende-se trazer uma discussão acerca de questões levantadas sobre a utilização desse método de avaliação em escala, mediante as análises dos dados comparativos.

2 Metodologia proposta

Para atingir o objetivo proposto neste trabalho vamos comparar os resultados das avaliações de tutores e alunos em um curso de pós graduação na modalidade presencial com o uso da estratégia *peer evaluation*.

2.1 Local e amostra do Estudo

O desenvolvimento da metodologia comparativa está apoiado na disponibilização de atividades dissertativas no ambiente virtual, previamente desenvolvidas por docentes, para uma amostra de 200 estudantes do curso de especialização em Gestão Estratégica (modalidade EaD) oferecida por um núcleo de pesquisa junto a uma das melhores universidades públicas do Brasil.

2.2 Ferramenta e Procedimentos

Para a realização deste estudo pretende-se utilizar o Workshop do Moodle, que é uma ferramenta livre, de código aberto e ideal para a aplicação deste tipo de metodologia, pois permite procedimentos avaliativos pelos tutores, por pares, autoavaliação e uma avaliação do tutor sobre as avaliações dos pares. A modelagem do trabalho seguirá praticamente quatro etapas descritas na figura 1.

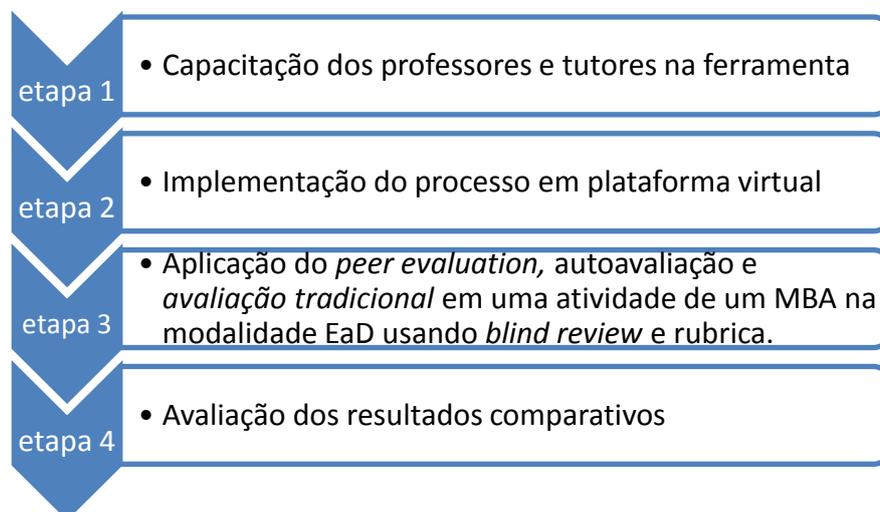


Figura 1 – Etapas para comparação da avaliação de tutores e alunos

1) Capacitação dos grupos (professores para confecção das rubricas e tutores e alunos para uso da ferramenta). Os professores foram orientados quanto a criação de rubricas que possibilitem aos alunos desenvolver um olhar crítico sobre a compreensão da pergunta, a coerência na resposta e o atendimento aos quesitos esperados para cada resposta. Nesta etapa os

alunos receberão treinamento quanto ao uso da ferramenta e serão orientados sobre a importância do estudo e a participação de cada um como agente avaliador de si mesmo de outros três colegas. Esta etapa já foi parcialmente concluída;

2) Implementação no ambiente virtual - Disponibilização e execução das atividades no ambiente virtual, com prazos para início e término. As atividades elaboradas pelos professores serão disponibilizadas no ambiente virtual e os ajustes nos parâmetros da ferramenta serão realizados para permitir que os alunos executem-nas em tempo hábil. Esta etapa está em processo de consolidação;

3) Aplicação da ferramenta - Período para correção das atividades e atribuição de notas, onde a configuração do Workshop também será delineada, garantindo que os alunos possam se autoavaliar e os pares acessem as atividades dos colegas (aleatoriamente no estilo *blind review*). Além disso, deve-se ressaltar que quando iniciada a etapa de avaliação, assim como para os tutores no processo tradicional, os alunos, agora vistos como agentes avaliadores, recebem uma rubrica para cada atividade. Esse gabarito, delineado pelo responsável (docente) pela disciplina, visa nortear os alunos tanto na autoavaliação quanto na avaliação dos colegas. Ainda para esta etapa é que serão feitos os ajustes na ferramenta, de modo que as atividades, de cada um dos 200 estudantes, sejam abertas para autoavaliação, encaminhadas aos respectivos tutores e encaminhadas aos pares, de modo que cada aluno tenha sua resposta avaliada por três colegas escolhidos aleatoriamente pelo sistema (nesse caso, cada aluno avaliará também as atividades de três outros colegas);

4) Análise dos resultados - Por fim uma avaliação geral dos dados e, possível, discrepância entre os três instrumentos de avaliação comparados (tutores, autoavaliação e pares) será realizada e testada por meio de análises estatísticas, de forma descritiva, com o intuito de corroborar estudos que apresentam a metodologia de *peer evaluation* como eficiente pedagogicamente [7],[8] e comprovar a hipótese de eficácia na avaliação em escala. A medida mais adequada para comparar as notas que dois avaliadores diferentes é o

consenso, encontrando a fração de notas atribuídas pelos alunos que coincidam com o próprio professor [9]. O consenso mede a exatidão de uma atribuição de nota pelo professor e a avaliação entre os pares. A metodologia, aqui empregada, pretende definir a correspondência estatística entre as atribuições dos pares e a dos tutores e professores. Esta medida não é simples de se obter, pois há diferentes interpretações sobre uma avaliação [10]. Em uma escala de 100 pontos, alguns pesquisadores consideram adequada uma diferença de ± 10 pontos e desvio-padrão de ± 1 [8].

2.3 Coleta de Dados e Apresentação

A coleta dos dados será feita pela equipe de autores deste trabalho, mediante chave de acesso ao sistema. Os dados coletados serão individuais e constituídos unicamente pelas notas atribuídas para cada aluno por autoavaliação, por pares e pelos tutores. As notas são dadas por cada agente e a média pode ser gerada, automaticamente, pelo sistema. Esse dado representará um bom indicador do desempenho geral do estudante, e mais, será importante na comparação das notas de cada agente em relação a média e os intervalos de confiança inferior (IC_{INF}) e superior (IC_{SUP}), que podem ser atribuídos pelo método *jackknife*.

A apresentação comparativa dos dados, com as possíveis discrepâncias e as análises estatísticas, entre as atribuições de notas de alunos e tutores poderá ser feita em gráficos de barras para valores inteiros e de dispersão para o desvio padrão (SD), além de tabelas comparativas entre os valores atribuídos pelos agentes, as médias e intervalos de confiança. Assim, espera-se ser possível comparar e avaliar, estatisticamente, se há consenso na atribuição de notas entre as metodologias de *peer evaluation* e a convencional, em função da amostra selecionada.

3 Resultados esperados

Este estudo permitira dar continuidade ao processo de validação da estratégia de *peer evaluation*. Espera-se que os resultados possam identificar

se esta estratégia é comparável a estratégia convencional de correção por parte dos tutores e a autoavaliação, além do prazo de *feedback* dispendido pelos avaliadores e a sua contribuição para a construção das competências sociais do aluno no aprendizado colaborativo.

Os resultados desta pesquisa podem viabilizar uma avaliação de qualidade em cursos onde existe um grande número de alunos.

Referências

- [1] Censo EAD.BR: **Relatório analítico da aprendizagem a distância no Brasil 2011** (2012), São Paulo: Pearson Education do Brasil.
- [2] Reja, U.; Manfreda, K.L., Hlbec, V.; Vehovar, V. **Open-ended vs. Close-ended Questions in Web Questionnaires**. <http://www.stat-d.si/mz/mz19/reja.pdf>
- [3] Black, Paul, William, Dylan, **Assessment in Education: Principles, Policy & Practice**, Mar1998, Vol. 5, Issue 1
- [4] Lee, H.-J., & Lim, C. (2012). Peer Evaluation in Blended Team Project-Based Learning: What Do Students Find Important?. **Educational Technology & Society**, 15 (4), 214–224.
- [5] Wever, B.D., Keer, H.V., Schellens, T, Valcke, M. (2011) Assessing Collaboration in a Wiki: The reliability of University Students' Peer Assessment. **The Internet and Higher Education**, 14(4), 201-206
- [6] Stevens, D.D., Levi, A.J. (2005) **Introduction to rubrics**, Sterling: Stylus Publishing, LLC
- [7] Ugulino, W., Marques, A.M., Pimentel, M., Siqueira S.W. (2009) **Avaliação Colaborativa: um Estudo com a Ferramenta Moodle Workshop**. XX Simpósio Brasileiro de Informática na Educação, Florianópolis - SC – 2009. ISSN: 2176-4301
- [8] Sadler, P.M., Good, E. . (2006). The Impact of Self- and Peer-Grading on Student Learning. **Educational Assessment**, 11(1), 1-31.
- [9] Burke, R.J. (1969). Some preliminary Data on the Use of Self-Evaluations and Peer Ratings in Assigning University Grades. **Journal of Educational Research**, 62, 444-448.

[10] Falchikov, N., Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. **Review of Educational Research**, 70(3), 287-322.