

DOI: 10.5748/9788599693124-13CONTECSI/PS-3891

SELF-, PEER- AND PROFESSOR ASSESSMENT USING RUBRICS IN ACCOUNTING

Lucas Almeida Fernandes Ribeiro - lucasalmeidafr@gmail.com

Jose Dutra Oliveira Neto - dutra@usp.br

Luiz Antonio Titton - titton@usp.br

Marcelo Ferreira da Silva - marcelofs@outlook.com

Additionally, the students' perception regarding the peer assessment approach was collected. A sample of 52 undergraduate accounting students enrolled in a research methodology class from a public university in Brazil. A tool called "workshop" from Moodle LMS were used during in-class activity of creating an academic . The student was required to post their own and do a self-assessment, using a rubric, on his/her own posted as well assess two other students' anonymously. An additional assessment was done by the instructor using the same rubric and compared to the students' assessment. The result shows that the self-assessment and peer-assessment are not statistically different, but they underrated themselves relative to the instructor. The results also show that student's perception of the importance of the peer assessment as a valuable tool for the learning process. The results are partially consistent with the literature. We conclude that the peer evaluation is a valuable tool for critical thinking and may help the education institution to assess the large number of assignments with small cost. The results of self and peer assessments do not demonstrate consistency with the results of instructor's assessment and may be due to the training of the students, quality of the rubric, and maturity differences of the students and instructors.

Keywords: Self-assessment, Peer-assessment, Teacher-assessment, rubrics

Introdução

Aprendizado ativo é o termo utilizado para definir uma abordagem de ensino construtivista que busca envolver os estudantes no processo de aprendizagem (Pritchard, 2008). Essa abordagem tem como foco o processo de aprendizagem e a forma como o estudante chega a determinado resultado, e não apenas o resultado em si (Walsh & Inala, 2010).

A teoria sobre esse método de ensino leva em conta o princípio básico de que as pessoas aprendem mais e retém melhor a informação quando são envolvidas e participam do processo de aprendizagem discutindo, questionando, buscando informações, realizando atividades práticas, do que quando apenas recebem a informação de forma teórica. Devido a isso a abordagem de aprendizado ativo é comumente definida como “aprender fazendo” (Gibbs, 2013).

No lugar das apresentações de fatos e informações por meio de palestras e aulas teóricas, a aprendizagem ativa ensina ao estudante como processar e utilizar o conhecimento obtido, promovendo o raciocínio, uma vez que permite ao estudante aplicar o que foi aprendido (Everly, 2013).

Conforme os estudantes progredem no seu processo de aprendizagem, os docentes devem ser capazes de avaliar essa evolução, enquanto os estimulam com atividades mais complexas. Pesquisadores vem encorajando a utilização de métodos de avaliação que incentivem a reflexão dos estudantes e promovam a colaboração com o processo de aprendizagem (Chen, 2010).

Nesse trabalho são abordados três métodos de avaliação, sendo eles: avaliação pelo professor (PROF), avaliação por pares (PEER) e aquela em que o estudante se auto avalia (SELF). Na avaliação PEER, os estudantes podem avaliar e receber avaliações de seus pares com ou sem o conhecimento de quem é a autoria do trabalho avaliado ou da avaliação recebida. Cabe ao professor escolher um ou mais métodos para adequar aos seus objetivos didáticos (Buzzetto-more & Alade, 2006; Chang, Tseng, & Lou, 2012).

Para que as avaliações ocorram de forma objetiva e o trabalho de apuração seja automatizado uma ferramenta utilizada no processo de avaliação é a rubrica. Essa ferramenta, além de viabilizar a implantação dos processos de avaliação mencionados, permite que haja maior agilidade no fornecimento de feedback aos estudantes.

A utilização de rubricas pode, como outras formas de auto avaliação, ser concebida distinguindo três abordagens no caso da SELF: conhecimento, aprendizagem cognitiva e resultados emocionais. A primeira abordagem refere-se ao conhecimento em si ou ao seu aumento no domínio objetivado. A segunda abordagem, da aprendizagem cognitiva, refere-se a compreensão sobre o conhecimento adquirido. Os resultados emocionais – terceira abordagem – é passível de três visões: reações, motivações e eficácia. Na visão das reações os estudantes refletem a satisfação com sua experiência educacional. A motivação é relativa ao grau de envolvimento deles para aplicar os conhecimentos adquiridos. A eficácia diz respeito a confiança deles com sua capacidade de executar as tarefas relacionadas com a formação recebida (Sitzmann, Brown, & Bauer, 2010; Sitzmann, Brown, Casper, Ely, & Zimmerman, 2008)

Fundamentalmente, existem dois tipos de rubricas: as avaliativas e as instrucionais. As primeiras são adequadas para quando avaliações de qualidade são necessárias sendo orientadas para dirigir as análises dos estudantes para avaliações mais objetivas. Por sua

vez, as rubricas instrucionais além dessa característica, oferecem a possibilidade de dar feedback sobre o que foi realizado de forma detalhada (Andrade, 2000a; Isaacson & Stacy, 2009; Moskal, 2000).

Simultaneamente, outra classificação para as rubricas é quanto a serem somativas ou formativas. As rubricas somativas baseiam-se em padrões, metas e critérios com a finalidade de medição e sem objetivo de oferecer retorno aos estudantes e, sendo assim, tem um caráter exclusivamente de mensuração. Nesse ponto de vista, as avaliações somativas são igualmente avaliativas, como na tipologia anterior. O que distingue a avaliação formativa, além de fornecer a possibilidade de dar feedback de forma detalhada, como as instrucionais, é a demonstração da existência de uma lacuna sobre a qual se deve dirigir os esforços de aprendizagem. O contexto em que o retorno é oferecido pressupõe alguma alteração para a redução da deficiência criando a possibilidade de que os estudantes desenvolvam a sua capacidade de avaliar o trabalho com base em um critério de padrão reconhecido de qualidade. Ao se adotar simultaneamente a avaliação SELF com a PEER, os estudantes realizam o processo de comparar as suas respostas com o padrão estabelecido e o de comparar as respostas dos outros estudantes com o padrão. Nessa dinâmica, a repetição da comparação com o padrão de qualidade estabelecido reforça a aprendizagem por meio da análise das deficiências observadas (Andrade, 2000b; Sadler-smith, 2007; Taras, 2005).

Tanto do ponto de vista do professor como dos estudantes, a avaliação PEER tem valor educativo relevante por poupar tempo e aumentar a aprendizagem do estudante. A percepção dos estudantes é a de que há impacto positivo na autoconfiança, comportamentos, identificação de pontos fortes e de limitações pessoais. Isso porque os pares contribuem adicionalmente para o desenvolvimento de competências e até mesmo podem alavancar o desenvolvimento de pensamento crítico a partir das contribuições cruzadas. Esse efeito dificilmente poderia ser obtido apenas com a auto avaliação. A avaliação PEER também contribui para melhorar as habilidades dos estudantes para relacionar os objetivos instrucionais na avaliação. O entendimento dos pontos fortes e fracos identifica as lacunas de conhecimento e promovem a compreensão e confiança melhorando o desempenho futuro (Bonk & Smith, 1998; Chang et al., 2012; Chen, 2010; Ng, 2014; Sadler & Good, 2006; Theising, Wu, & Heck Sheehan, 2014).

As mensurações confiáveis de performance podem ser melhoradas por meio de rubricas especialmente se forem analíticas, específicas para um tópico e complementadas com exemplos e treinamento no uso delas. Entretanto, as rubricas não promovem julgamentos de performance por si mesmas, elas necessitam de um quadro abrangente de validação. A razão do potencial educacional das rubricas é o fato que tornam explícitas as expectativas e os critérios, e por isso facilitam a auto avaliação e o feedback (Jonsson & Svingby, 2007).

Esse conjunto contribui decisivamente para o desenvolvimento de pensamento crítico (*critical thinking*). Essa abordagem inclui uma proposta-mestre para a aplicação, evidenciando a compreensão e interiorização dessa competência especialmente concebida para medir o desenvolvimento do pensamento crítico, mais do que avaliar a realização de tarefas (Paul & Elder, 2007).

A implementação desse método de avaliação fornece um *feedback* contínuo aos estudantes durante o processo de aprendizagem, ajuda a desenvolver o pensamento crítico e a capacidade de julgamento por meio das avaliações (Hassan, Fox & Hannah, 2014).

Durante a avaliação entre os pares, os estudantes podem comparar seu trabalho com os apresentados nas rubricas reforçando a reflexão e o pensamento aprofundado (Tsivitanidou & Constantinou, 2015).

O projeto de rubricas deve medir as competências de acordo com a dimensão que se pretende medir, que pode ser a do conhecimento, da compreensão ou de capacidades de nível superior. Nesse estudo foi utilizada nas rubricas a dimensão dos elementos do pensamento crítico associado às tarefas desenvolvidas pelos estudantes no sentido da aprendizagem cognitiva em que o objeto a ser medido é o conhecimento baseado em fatos e tarefas de conhecimento.

Uma pesquisa realizada com estudantes que trabalham sobre o mesmo assunto em sala de aula mostrou evidências de diferenças entre as avaliações SELF, PEER e PROF. Verificou-se que SELF e PROF são consistentes entre si, mas PEER mostrou diferenças significativas com os outros dois métodos (Chang, Liang, & Chen, 2013; Chang et al., 2012). Um estudo posterior testou uma nova configuração utilizando o método baseado em internet e descobriu que neste novo ambiente os três métodos convergiram para os mesmos resultados (Chang et al., 2012). Nestes dois estudos, as avaliações pelos pares não foram anônimas.

O ambiente colaborativo pode levantar questões sobre a avaliação dos colegas especificamente em termos de anonimato (Chen, 2010). Além disso, os estudantes mostraram-se divididos sobre o impacto do anonimato e amizade (Theising et al., 2014). Os estudantes podem se sentir desconfortáveis em criticar o desempenho de cada outro (Chen, 2010). Observou-se que, mesmo entre os diferentes grupos, os estudantes avaliam os seus próprios membros do grupo e outros, sem restrições sobre quem era o outro grupo a ser avaliado, a tendência para pesar a avaliação positiva de seu próprio grupo sobre o outro é um fato (Sadler & Good, 2006). Pode-se esperar um processo de avaliação mais justo quando os pares percebem o ambiente de aprendizagem como seguro e menos propenso a pressão dos colegas (van Gennip, Segers & Tillema, 2010). A configuração do anonimato em cursos on-line pode definir um contexto que afeta positiva ou negativamente a confiança na avaliação. Neste contexto particular, verificou-se que nas avaliações anônimas por pares os estudantes apresentaram um melhor desempenho nas tarefas que envolvem a escrita e forneceram feedback mais crítica aos seus pares do que os estudantes que participam de avaliações onde a identidade dos avaliadores e avaliados são reveladas (Lu & Bol, 2007).

Estudos evidenciam que a avaliação dos professores tende a ser mais criteriosa e crítica, atribuindo notas inferiores quando comparada àquelas atribuídas pelos estudantes a si mesmos. Da mesma forma, as auto avaliações tendem a ser mais generosas do que aquelas atribuídas aos seus pares (Lin, Liu, & Yuan, 2001; Sadler & Good, 2006).

Entretanto, resultados obtidos por meio da auto avaliação e da avaliação por pares comparativamente aqueles encontrados por meio da avaliação do professor apontam haver consistência entre os métodos em alguns aspectos (Chang, Tseng & Lou, 2012).

Embora as pesquisas sobre esses métodos de avaliação abranjam inúmeras disciplinas, no ensino da contabilidade esta área de estudo continua a ser um tema relevante e objeto de pesquisa em diversos trabalhos (Freeman, 1995; Hassan, Fox & Hannah, 2014; Kilpatric, Linville & Stout, 2001; Larres, Ballantine & Whittington, 2003; Lirely et al., 2011). No entanto, cabe ressaltar que ainda não existe convergência nos resultados obtidos, o que evidencia a necessidade de novos estudos sobre o tema.

No estudo de Hassan, Fox & Hannah (2014), que comparou a auto avaliação (SELF), avaliação por pares (PEER) e a avaliação pelo professor (PROF) utilizando como amostra dois grupos de estudantes de contabilidade, os resultados apresentados demonstraram que existe uma diferença significativa entre os métodos de avaliação, indicando que SELF e PEER não são precisas quando comparadas com PROF.

Os resultados da pesquisa de Freeman (1995), que estudou a precisão de PEER comparando as diferenças entre PEER e PROF de estudantes do último ano de um curso de graduação em contabilidade e finanças, indicaram que as notas obtidas com os dois métodos de avaliação não são significativamente diferentes, embora individualmente apresentem variações consideráveis.

Outro estudo sobre auto avaliação, realizado com estudantes do primeiro ano do curso de graduação em contabilidade, demonstrou que os resultados obtidos com esse método foram imprecisos e apontou que existe uma grande diferença entre os estudantes do primeiro ano, sendo que alguns estudantes são capazes de realizar a auto avaliação com mais precisão (Larres, Ballantine & Whittington, 2003).

Esse estudo visa **apurar a convergência entre os três tipos de avaliação, auto avaliação (SELF), avaliação por pares (PEER) e a do professor (PROF) por meio de rubrica em disciplina de metodologia no curso de Ciências Contábeis.**

Esse estudo visa responder as seguintes questões de pesquisa:

- 1) As notas dos avaliadores SELF, PROF e PEER são estatisticamente diferentes? Qual é a maior e a menor?
- 2) As notas por gênero são estatisticamente diferentes? Qual é a maior e a menor?
- 3) Qual a percepção de satisfação de estudante com relação a auto avaliação e avaliação de pares?

Metodologia

Amostra de participantes

Os participantes da pesquisa foram 52 estudantes de duas disciplinas sobre pesquisa acadêmica de um curso superior em contabilidade de uma faculdade pública no Brasil. Esta amostra foi composta de 19 estudantes do sexo feminino (36,5%) e 33 do sexo masculino (63,5%).

Sistema utilizado

O ambiente virtual de aprendizagem utilizado foi o Moodle®, por meio da ferramenta denominada “workshop”. A escolha de dois avaliados por estudante foi feita de forma aleatória pelo sistema Moodle®. Os estudantes avaliam a sua própria atividade, bem como a atividade de outros dois colegas, sem se identificar, utilizando um critério previamente concebido para avaliar a tarefa. Isto dá ao estudante o *feedback* sobre o que é esperado com o potencial de promover ou melhorar a sua aprendizagem (Jonsson & Svingby, 2007).

Desenvolvimento da rubrica de avaliação

As rubricas para a avaliação do resumo, a partir de um artigo científico, foram criadas usando seis dimensões sendo elas: Contexto; Gap; Objetivo; Material e métodos; e

Resultados, sendo que cada dimensão tem uma nota que pode variar de um a quatro. A proposta é avaliar se o estudante demonstrou conhecimento e compreensão sobre o conteúdo do artigo acadêmico a um nível onde o estudante pode produzir um resumo acerca do pensamento do autor. A pontuação mínima é de 6 e a máxima é de 24.

Para cada dimensão, existe um conjunto de respostas possíveis que serão escolhidas pelo avaliador e devem representar o domínio do estudante em relação à dimensão. A matriz de avaliação tem 6 linhas (dimensões) e 4 colunas (respostas possíveis dos estudantes em cada dimensão). O estudante deve utilizar o mouse para escolher uma resposta para cada dimensão que mais se aproxima da realidade. O estudante avaliador visualiza na matriz de avaliação um padrão de referência desejado para as respostas dos avaliados. A apresentação de uma referência para a avaliação aumenta em alguns casos a propensão de feedbacks válidos e confiáveis (Jones & Alcock, 2014).

A rubrica completa, com o cabeçalho representativo do padrão de qualidade da resposta desejada está descrito nas tarefas (T) de 1 a 6 (Apêndice 1).

Procedimentos



Figura 1 – Procedimentos

Para a atingir o objetivo proposto, foi definido um procedimento descrito na Figura 1 que envolve a capacitação do estudante na ferramenta, criação do resumo, avaliação e percepção do estudante em relação a abordagem adotada.

Capacitação prévia – Treinamento

Foi realizado um treinamento utilizando a ferramenta “workshop” com um artigo teste da área de contabilidade. Esta etapa é fundamental para que o estudante entenda a ferramenta e os conceitos utilizados.

Etapa 1 - A tarefa de criação do resumo

Cada estudante recebeu uma cópia em papel (embora a mesma estivesse disponível online) do mesmo artigo acadêmico e teve a tarefa de escrever as seis seções de um resumo

padrão. Cada seção foi associada a uma cor para facilitar a compreensão e correção. As seções são Contexto (amarelo), Lacuna ou gap (azul), Objetivo (rosa), Material e métodos (cinza), Resultados (verde) e Conclusões (vermelho). No final da aula, o sistema “workshop Moodle®” foi reposicionado para a fase de avaliação, onde os estudantes podem avaliar o trabalho desenvolvido pelos seus pares e os seus próprios trabalhos. Os estudantes deveriam postar os seus resumos, de forma individual, até o final da aula utilizando a ferramenta *workshop* do Moodle®.

Etapa 2 - A tarefa de avaliação

Cada participante deveria avaliar outros 2 estudantes utilizando a rubrica. Os estudantes são escolhidos de forma aleatória e anônima pelo Moodle®. O instrutor também realizou a mesma avaliação usando a mesma matriz de avaliação (rubrica). Ao final, obteve-se até 3 avaliações por estudante.

Etapa 3 – Percepção do estudante em relação a avaliação de pares

Para a apuração na dimensão emocional (satisfação) foi utilizado questionário em formulário anônimo (Apêndice 2). O questionário estava hospedado em um servidor comercial, o que permitia o anonimato aos estudantes.

Resultados e discussão

Para responder a primeira questão de pesquisa, se as notas dos 3 avaliadores SELF, PROF e PEER são estatisticamente diferentes, foi utilizado o Método da Análise de Variância (ANOVA) de Medidas Repetidas, além da estatística descritiva (Tabela 1). Como resultado obtivemos que $-F(2;120) = 6,03$ $p = 0,003^*$, ou seja, há diferença significativa entre as notas dos 3 avaliadores.

Estatística Descritiva			
	Média	Desvio Padrão	N
Nota SELF	4,9590	1,11913	61
Nota_PROF	5,5082	1,09351	61
Nota_PEER	4,9721	1,16077	61

Tabela 1 - Comparação entre as notas Self, Prof e Peer

Conclui-se que as notas do professor são estatisticamente diferentes das atribuídas por meio da auto avaliação (SELF) e da realizada pelos pares (PEER). No entanto, nota-se que os resultados obtidos por meio de SELF e PEER são próximos e apresentam uma convergência.

Para responder a segunda questão de pesquisa sobre qual é a maior e qual é a menor, pelo teste *post hoc* de Bonferroni temos que a nota do Professor é significativamente superior à do estudante ($p = 0,006$) e à média do Peer ($p = 0,02$).

Para uma análise mais profunda das diferenças entre as notas do professor, auto avaliações e avaliações por pares foi calculado o percentual de quanto as notas do professor estão superiores às notas dos estudantes PEER e SELF Esta análise tem como

objetivo identificar se a correlação é suficiente para estabelecer uma relação a partir da qual se apure a nota do professor a partir da nota do estudante.

Inicialmente avaliamos por meio de média e variância dos percentuais das notas do professor sobre as notas dos estudantes (PEER e SELF), na Tabela 2- Comparação entre as notas SELF com PROF

	N	Minimo	Maximo	Média	Desvio Padrão
P	61	-50,00	75,00	-7,0245	25,24756
Válido (listwise)	N 61				

Tabela 2- Comparação entre as notas SELF com PROF

Com base nessa primeira abordagem, pode-se calcular que a nota do estudante deveria ser majorada em 7,03% para se igualar àquela do professor. A amplitude entre máximo e mínimo vai de 50 a 75%, e por isso, essa informação serve apenas para confirmar que a nota do professor é superior àquela dada pelos estudantes, sem que isso signifique ser possível um cálculo preciso a esse respeito.

Para uma análise mais profunda das diferenças entre as notas do professor, auto avaliações e avaliações dos pares, foi utilizado também o método de regressão linear. Primeiro foi realizada a regressão entre as notas da auto avaliação (SELF) com a nota dada pelo professor (PROF) para avaliar se existe alguma correlação entre essas duas variáveis e as notas de pares (PEER) com a nota dada pelo professor (PROF).

		Nota_Prof
Nota_SELF	Correlação Pearson	,293*
	Sig. (bicaudal)	,022
	N	61
Nota_Peer	Correlação Pearson	,168
	Sig. (bicaudal)	,197
	N	61
Média_PEER_e_SELF	Correlação Pearson	,294*
	Sig. (bicaudal)	,021
	N	61

*. Correlação significativa para o nível 0.05 (bicaudal).

Tabela 3 - Correlação entre as notas dos estudantes e a do professor

Com base nos dados da Tabela 3 - Correlação entre as notas dos estudantes e a do professor

, é possível afirmar que a correlação entre a nota SELF e a PROF é baixa, apenas 0,29, para que seja possível atribuir uma em função da outra. Os dados apresentados na Figura 2 demonstram essa baixa correlação:

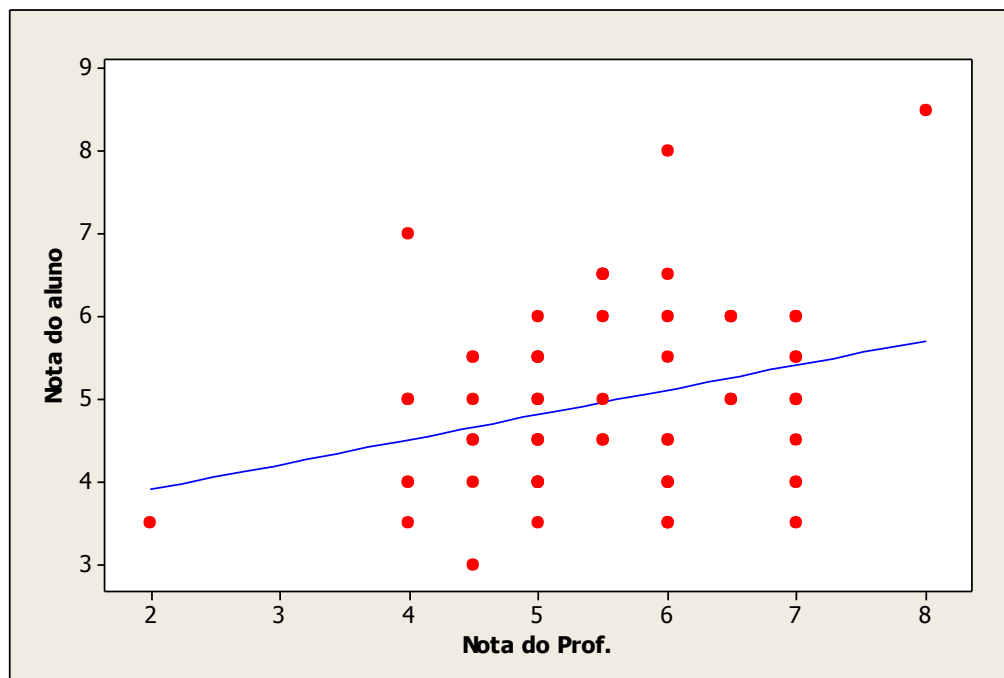


Figura 2 - Plotagem das notas PROF com as notas SELF

Observa-se no gráfico que para cada nota PROF há inúmeras notas SELF, confirmando graficamente o observado na tabela de regressão.

Esse mesmo contexto repete-se quando se observa na Figura 3, as notas PROF comparadas com notas PEER:

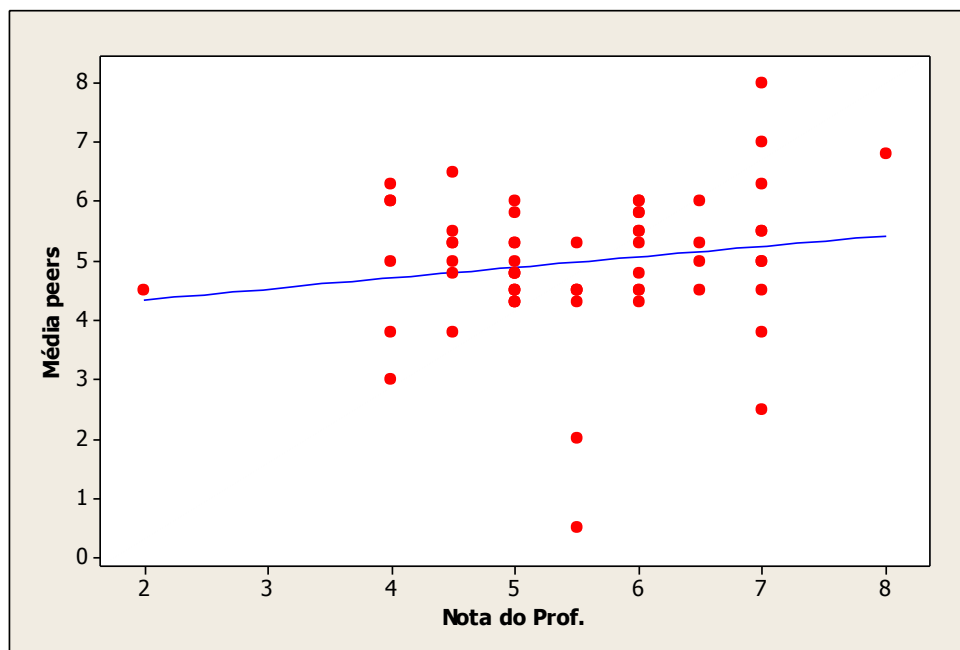


Figura 3- Plotagem das notas PROF com as notas PEER

Esses dados apontam para uma diferença entre o que existe em pesquisas anteriores, em especial ao trabalho realizado por Chang et al 2012 e Chang et al 2013 em que SELF e PROF demonstraram consistência entre si com diferenças significativas quanto ao PEER.

No trabalho de Chang et al. (2012) o nível educacional dos estudantes, o ambiente de avaliação e os procedimentos empregados são apontados como prováveis motivos para a divergência nos resultados. Não obstante, fatores como o treinamento empregado para as avaliações e as rubricas utilizadas também podem ter influenciado no resultado alcançado.

Comparando com resultados obtidos com trabalhos realizados em disciplinas do curso de contabilidade podem-se verificar certas similaridades. Nos trabalhos de Hassan, Fox & Hannah (2014), Larres, Ballantines & Whittington (2003) e Lirely et al. (2011) não foi encontrada convergência entre os métodos de avaliação utilizados.

Entretanto, o estudo de Freeman (1995) apontou não existir diferenças significativas entre a avaliação dos pares e a do instrutor afirmando que as notas obtidas foram parecidas.

Para responder a segunda questão de pesquisa considerando gênero, foi realizado o teste t além da estatística descritiva com relação a gênero.

Os resultados para o teste t foram $T = 0,79$ $p = 0,43$ que indicam que não há diferença significativa.

	Sexo	N	Média	Desvio padrão
Auto_Nota	M	35	5,0571	1,10992
	F	26	4,8269	1,13967

Tabela 4 - Os resultados para estatística descritiva média e desvio padrão

Na nossa amostra, não foi detectada diferença significativa entre os gêneros.

Para responder à questão de pesquisa sobre a satisfação do estudante sobre o novo processo de avaliação foi desenvolvido um instrumento de forma anônima, hospedado em site comercial diferente do ambiente virtual de aprendizagem, de 17 questões de múltipla escolha que foram aplicadas ao final do curso.

A média de estudantes que discordam totalmente ou discordam muito foi de 3,2 respostas, enquanto que a média de discordo ou não concordo nem concordo somam 15,9. As respostas superiores (concordo, concordo muito, concordo totalmente) somam 32,8. Ou seja, do total de 52 respostas possíveis, 63% dos estudantes atribuem boas avaliações com relação a auto avaliação e avaliação por pares (Figura 4).

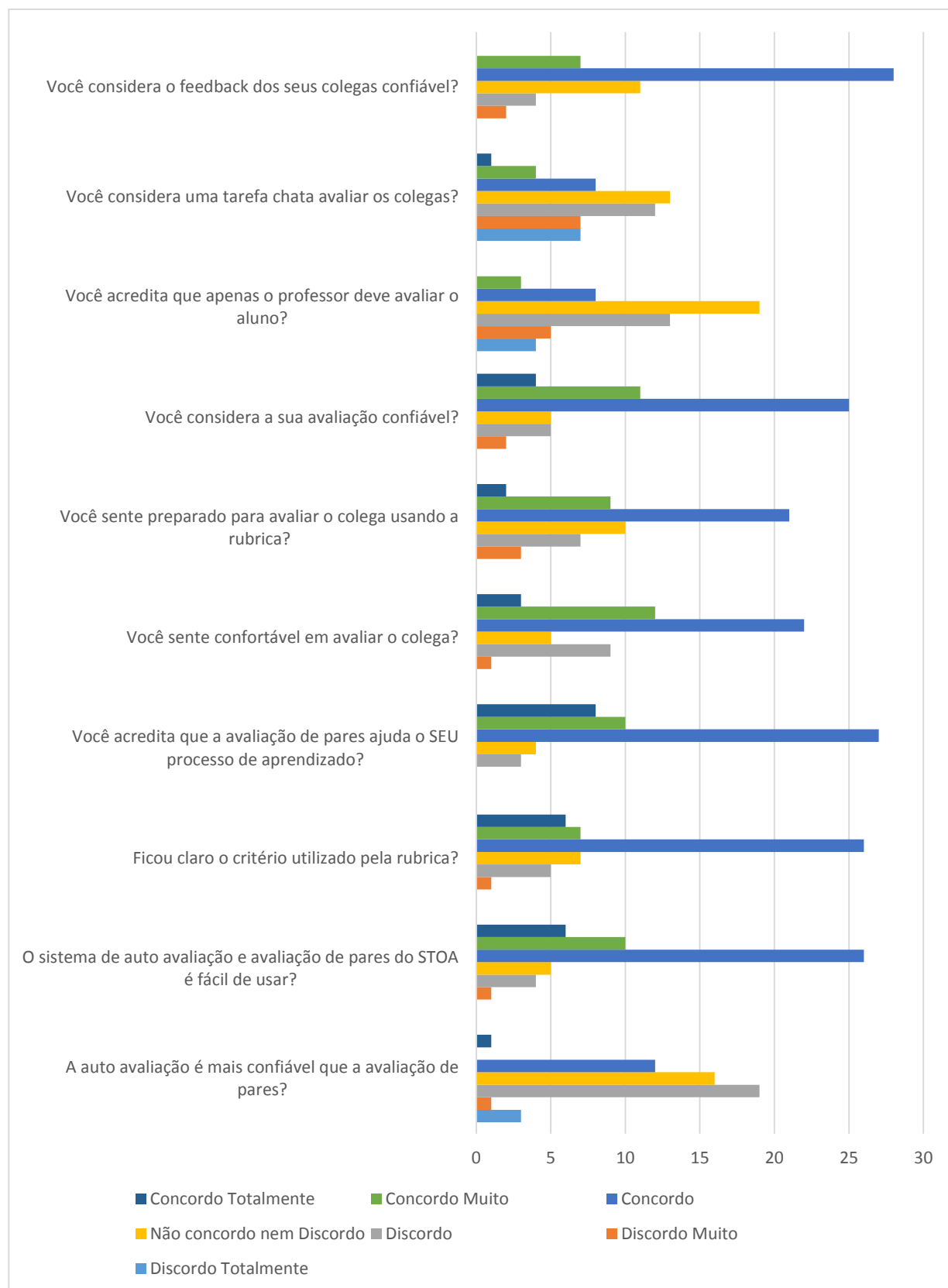


Figura 4 - Resultado da avaliação de satisfação

Esses resultados corroboram as pesquisas anteriores, não apenas quanto a satisfação, mas também quanto a motivação, como já identificado por Sitzmann, et al (2008).

Conclusão

Pode-se concluir que o modelo de uso de rubricas para a avaliação de estudantes de graduação em contabilidade mostrou-se bastante produtivo, principalmente pela forma como o estudante faz contato com a resposta considerada didaticamente adequada para cada dimensão. Esse processo desenvolve o pensamento crítico e uma capacidade de nível superior ao dar a chance de avaliar o seu próprio trabalho e o de outros. A repetição do processo alavanca efetivamente os resultados.

Por outro lado, as diferenças encontradas entre a auto avaliação e a avaliação dos pares quando comparadas com a avaliação do professor demonstraram não haver consistência entre os métodos analisados. A falta de convergência entre os resultados obtidos por meio de SELF e PEER com os resultados da avaliação PROF deixou margem para questionamentos sobre outros fatores que poderiam estar afetando a precisão das notas obtidas com esses métodos de avaliação.

Entre os fatores aventados ao longo das análises, sugere-se que a qualidade da rubrica, a formação dos alunos e a diferença na rigidez e na maturidade de estudantes e professor podem ter afetado os resultados.

O clima social e organizacional da sala de aula poderia, também, afetar a forma como as avaliações PEER foram realizadas, mesmo com a utilização do modelo de avaliações anônimas. Outro aspecto possível que pode ter influenciado é o de que diferentes professores e sua postura ao longo do tempo quanto a rigidez em avaliações poderia resultar em distanciamento quanto às avaliações PEER.

Mesmo assim, as consequências dessa pesquisa para o estudante demonstraram que o conjunto formado pelo processo do uso de rubricas com os três métodos de avaliação acrescenta ganhos na aprendizagem, o que é muito importante para o estudante. Do ponto de vista da instituição de ensino, há indícios de que esse ganho de produtividade na aprendizagem acontece com o mesmo custo ou redução, se considerarmos o potencial para a aplicação em escala nos ambientes de aprendizagem em massa. O trabalho do professor é pouco alterado em função desse processo. Para instituições com grande número de estudantes surge a possibilidade de avaliar um grande número de estudantes por um custo reduzido. Outro fator positivo diz respeito às dúvidas expostas a partir dos estudantes, que tornam-se mais pontuais e dirigidas para pontos focais do conteúdo pela forma como o andamento das atividades acontece – o que favorece e valoriza seu trabalho. O fato de não haver diferenças entre os gêneros pode significar que ambos se beneficiam igualmente dos resultados.

Os resultados da percepção dos estudantes em relação a avaliação, evidenciam que os estudantes estão dispostos a aceitar novos métodos de avaliação que favoreçam o seu aprendizado.

A concepção, o desenvolvimento e a validação da rubrica são fatores importantes no processo de avaliação e sua utilização na avaliação favorece a focalização de aspectos granulares do conteúdo e dá a oportunidade da repetição do contato com o conteúdo por meio do acesso frequente e com visualização dos padrões desejados de qualidade.

Referências

- Andrade, H. G. (2000a). Using Rubrics to Promote Thinking and Learning. *Educational Leadership*, 57(5), 13–18.
- Andrade, H. G. (2000b). Using Rubrics to Promote Thinking and Learning. *Educational Leadership*, 57(5), 13–18.
- Bonk, C. J., & Smith, G. S. (1998). Alternative instructional strategies for creative and critical thinking in the accounting curriculum. *Journal of Accounting Education*, 16(2), 261–293.
- Buzzetto-more, N. A., & Alade, A. J. (2006). Best Practices in e-Assessment. *Journal of Information Technology*, 5, 251–269.
- Chang, C.-C., Liang, C., & Chen, Y.-H. (2013). Is learner self-assessment reliable and valid in a Web-based portfolio environment for high school students? *Computers & Education*, 60(1), 325–334. <http://doi.org/10.1016/j.compedu.2012.05.012>
- Chang, C.-C., Tseng, K.-H., & Lou, S.-J. (2012). A comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a Web-based portfolio assessment environment for high school students. *Computers & Education*, 58(1), 303–320. <http://doi.org/10.1016/j.compedu.2011.08.005>
- Chen, C. (2010). The implementation and evaluation of a mobile self- and peer-assessment system. *Computers & Education*, 55(1), 229–236. <http://doi.org/10.1016/j.compedu.2010.01.008>
- Everly, M. C. (2013). Are students' impressions of improved learning through active learning methods reflected by improved test scores. *Nurse Education Today*, 33(1), 148–151.
- Gibbs, G. Learning by Doing. Oxford: Oxford Centre for Staff and Learning Development, 2013.
- Freeman, M. (1995). Peer Assessment by Groups of Groups Work. *Assessment & Evaluation in Higher Education*, 20(3), 289–300.
- Hassan, O. A. G., Fox, A., & Hannah, G. (2014). Self- and Peer-Assessment: Evidence from the Accounting and Finance Discipline. *Accounting Education*, 23(3), 225–243.
- Isaacson, J. J., & Stacy, A. S. (2009). Rubrics for clinical evaluation : Objectifying the subjective experience. *Nurse Education in Practice*, 9(2), 134–140. <http://doi.org/10.1016/j.nepr.2008.10.015>
- Jones, I., & Alcock, L. (2014). Peer assessment without assessment criteria. *Studies in Higher Education*, 39(10), 1774–1787.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics : Reliability , validity and educational consequences, 2, 130–144. <http://doi.org/10.1016/j.edurev.2007.05.002>
- Kilpatrick, D. J., Linville, M. & Stout, D. E. (2001). Procedural justice and the development and use of peer evaluations in business and accounting classes. *Journal of Accounting Education*, 19, 225–246.

- Larres, P. M., Ballantine, J. & Whittington M. (2003). Evaluating the validity of self-assessment: measuring computer literacy among entry-level undergraduates within accounting degree programmes at two UK universities. *Accounting Education*, 12(2), 97–112.
- Lirely, R., Keech, K. M., Vanhook, C. & Little, P. (2011). Developmental and evaluative contextual usage of peer assessment of research presentations in a graduate tax accounting course. *International Journal of Business and Social Science*, 2(23), 89–94.
- Lin, S. S. J., Liu, E. Z. F., & Yuan, S. M. (2001). Web-based peer assessment : feedback for students with various thinking-styles, (January), 420–432.
- Lu, R., & Bol, L. (2007). A Comparison of Anonymous Versus Identifiable e-Peer Review on College Student Writing Performance and the Extent of Critical Feedback. *Journal of Interactive Online Learning*, 6(2), 100–115.
- Moskal, B. M. (2000). Scoring Rubrics : What , When and How ? *Practical Assessment, Research & Evaluation*, 7(3), 1–7. Retrieved from <http://pareonline.net/getvn.asp?v=7&n=3>
- Ng, E. M. W. (2014). Using a mixed research method to evaluate the effectiveness of formative assessment in supporting student teachers' wiki authoring. *Computers & Education*, 73, 141–148. <http://doi.org/10.1016/j.compedu.2013.12.016>
- Paul, R., & Elder, L. (2007). *CriticalThinking.pdf* (Special Ed). Berkeley, USA: The Foundation for Critical Thinking. Retrieved from www.criticalthinking.org
- Pritchard, A. Ways of Learning: learning theories and learning styles in the classroom (Second Edition). New York: Taylor & Francis e-Library, 2009.
- Sadler, P. M., & Good, E. (2006). The Impact of Self- and Peer-Grading on Student Learning. *Educational Assessment*, 11(1), 1–31. http://doi.org/10.1207/s15326977ea1101_1
- Sadler-smith, E. (2007). in *Management Education*, 6(2), 186–205.
- Sitzmann, T., Brown, K. G., & Bauer, K. N. (2010). Self-Assessment of Knowledge : A Cognitive Learning or Affective Measure ?, 9(2), 169–191.
- Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K., & Zimmerman, R. D. (2008). A review and meta-analysis of the nomological network of trainee reactions. *The Journal of Applied Psychology*, 93(2), 280–95. <http://doi.org/10.1037/0021-9010.93.2.280>
- Taras, M. (2005). Assessment – Summative and Formative – Some Theoretical Reflections. *British Journal of Educational Studies*, 53(4), 466–478.
- Theising, K., Wu, K., & Heck Sheehan, A. (2014). Impact of peer assessment on student pharmacists' behaviors and self-confidence. *Currents in Pharmacy Teaching and Learning*, 6(1), 10–14. <http://doi.org/10.1016/j.cptl.2013.09.020>
- Tsivitanidou, O. E., & Constantinou, C. P. (2015). A study of students' heuristics and strategy patterns in web-based reciprocal peer assessment for science learning. *Internet and Higher Education*, 29, 12–22.
- van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2010). Peer assessment as a

collaborative learning activity: The role of interpersonal variables and conceptions.
Learning and Instruction, 20, 280–290.

Walsh, A., & Inala, P. Active Learning techniques for Librarians. Oxford: Chandos Publishing, 2010.