



UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Métricas objetivas e critérios para apoiar a
avaliação da participação em bate-papo educacional

Marcelo da Costa Zeferino

Orientador
Mariano Pimentel

RIO DE JANEIRO, RJ – BRASIL
SETEMBRO DE 2014

Métricas objetivas e critérios para apoiar a avaliação da participação
em bate-papo educacional

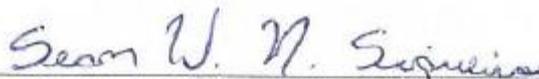
Marcelo da Costa Zeferino

DISSERTAÇÃO APRESENTADA COMO REQUISITO PARCIAL PARA
OBTENÇÃO DO TÍTULO DE MESTRE PELO PROGRAMA DE PÓS-
GRADUAÇÃO EM INFORMÁTICA DA UNIVERSIDADE FEDERAL DO ESTADO
DO RIO DE JANEIRO (UNIRIO). APROVADA PELA COMISSÃO
EXAMINADORA ABAIXO ASSINADA.

Aprovada por:



Mariano Pimentel, DSc (Orientador)
Universidade Federal do Estado do Rio de Janeiro – UNIRIO



Sean Wolfgang Matsui Siqueira, DSc
Universidade Federal do Estado do Rio de Janeiro – UNIRIO



Edméa Oliveira dos Santos, DSc
Universidade do Estado do Rio de Janeiro - UERJ

RIO DE JANEIRO, RJ – BRASIL
SETEMBRO DE 2014

Z43 Zeferino, Marcelo da Costa.
Métricas objetivas e critérios para apoiar a avaliação da participação em bate-papo educacional / Marcelo da Costa Zeferino, 2014.
109 f. ; 30 cm

Orientador: Mariano Pimentel.
Dissertação (Mestrado em Informática) - Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2013.

1. Grupos de bate-papo pela Internet. 2. Ensino via web. 3. Ensino à distância. 4. Estudantes - Avaliação. I. Pimentel, Mariano.
II. Universidade Federal do Estado do Rio de Janeiro. Centro de Ciências Exatas e Tecnológicas. Curso de Mestrado em Informática.
III. Título.

CDD - 005.5

Dedicatória

Dedico esse trabalho primeiramente aos meus pais, que me ensinaram a importância do esforço para conseguir alcançar nossos objetivos nesta vida e à minha família, que me proporciona a base necessária para alcançar meus objetivos.

Agradecimentos

Agradeço a Deus por ter me guiado e proporcionado a força necessária em todos os momentos onde as dificuldades se apresentaram como barreira para meu avanço. Agradeço por ter conseguindo vencer todas essas barreiras.

A minha esposa e filhos, que me apoiaram nos últimos anos, compreendendo os momentos onde foi necessária dedicação extra para cumprir com as responsabilidades diversas. A meu filho Pedro, que me ensinou e me ensina sobre muitas coisas a cada dia, principalmente sobre mim mesmo.

Ao meu orientador, Mariano Pimentel, pela dedicação, apoio e confiança em minha capacidade de seguir em frente. O cuidado e a responsabilidade que demonstrou com a construção do meu conhecimento e com os trabalhos que foram necessários me ensinou uma nova perspectiva sobre o resultado do meu esforço.

Agradeço também aos professores do PPGI que contribuíram para minha formação e a meus amigos do Grupo Comunicatec, que me auxiliaram a alcançar meus objetivos.

.

ZEFERINO, Marcelo da Costa. Métricas objetivas e critérios para apoiar a avaliação da participação em bate-papo educacional. UNIRIO, 2014. [106] páginas. Dissertação de Mestrado. Departamento de Informática Aplicada, UNIRIO.

RESUMO

O principal insumo para a avaliação da participação em sessões de bate-papo são as mensagens enviadas pelos alunos, pois representam o histórico da comunicação. A análise humana e manual de todas as mensagens enviadas demanda tempo dos educadores e pode dificultar o trabalho de avaliação da participação na sessão. Nesta dissertação foram estudadas métricas objetivas e critérios visando a definição de um método para qualificação da participação dos alunos em uma sessão de bate-papo educacional, com o objetivo de apoiar educadores que utilizem o bate-papo na Educação *Online*. O trabalho apresenta uma proposta de métricas e critérios com diferentes visões da participação, que permitiram a geração de um Indicador da Participação para cada aluno, apresentado em relatório que classifica participação dos alunos na sessão. A solução foi aplicada em estudo de caso exploratório com uma turma de graduação em informática, onde foi possível avaliar a aceitação da solução proposta e refiná-la com base na discussão dos resultados obtidos. As notas atribuídas pelo método proposto para a qualificação da participação foram comparadas com as notas do tutor da sessão e com a autoavaliação dos alunos, com o objetivo de medir a proximidade entre as notas. Foram realizados ajustes que indicam que a nota atribuída pelo método proposto nesta pesquisa foi próxima da nota atribuída manualmente pelo tutor (com análise humana das mensagens).

Palavras-chave: Educação *Online*, Educação a Distância, Bate-papo, Avaliação em Bate-papo, Avaliação da Participação, Participação em Sessões de Bate-papo.

ABSTRACT

The main input for the participation evaluation in chat sessions are the messages sent by the users, because they represent the history of communication. The human and manual analysis of all messages sent by the users expends time from the educators and this work can difficult the evaluation of participation in the session. On this dissertation were studied objective metrics and rules to define a method for qualify the participation of students on educational chat sessions, in order to support educators that uses chat on Online Education. The presented research proposes a set of metrics and rules with different visions for participation, allowing the generation of an indicator for Participation Performance for each student, presented in a report that classifies the participation of each student on a chat session. The proposed solution was applied on an exploratory study case with students, allowing the evaluation of the proposed solution and the refinement with the feedback and results obtained. The results of participation analysis proposed on this dissertation was compared with the educator's assessment and the student's self-assessment, aiming to measure the proximity between results. The adjustments indicate that the results of the analysis given by the proposed method was similar with the results of analysis executed manually by the educator (with human analysis of each messages).

Keywords: Online Education, Distance Education, Chat, Chat Evaluation, Participation Evaluation, Participation in Chat Sessions.

Sumário

1.	Introdução	9
1.1	Bate-papo para Educação Online	9
1.2	Objetivo da pesquisa.....	12
1.3	Visão Geral da Pesquisa	13
1.3.1	Apresentação do Problema.....	13
1.3.2	Proposta de Solução	13
1.3.3	Avaliação da Solução.....	14
1.4	Organização do texto da Dissertação.....	15
2.	Avaliação da Participação em Bate-papo.....	16
2.1	Avaliação na Educação.....	16
2.2	Avaliação na Educação online.....	18
2.3	Avaliação da Participação na Educação Online	21
2.4	Avaliação da Participação em sessões de Bate-Papo educacional	21
3.	Problematizando a Avaliação da Participação em Bate-Papo.....	27
3.1.	Dificuldades na avaliação da participação em sessões de bate-papo	27
3.2.	A análise humana de mensagens enviadas durante a sessão de bate-papo consome muito tempo dos educadores	28
3.3.	Dificuldades para a análise automática (ou semiautomática) das mensagens..	29
4.	Proposta de Solução: Métricas Objetivas e Critérios para apoiar a Avaliação da Participação em sessão de bate-papo educacional.....	31
4.1.	Métrica: Tempo de Presença na Sessão.....	32
4.2.	Métrica: Quantidade de Mensagens.....	32
4.3.	Métrica: Quantidade Média de Caracteres por Mensagem.....	35
4.4.	Métrica: Pertinência das Mensagens	37
4.5.	Nota Final – Indicador da Participação	38
5.	Avaliação das Métricas e Critérios: um estudo de caso em uma sessão real de bate-papo.....	40
5.1	Plano e contexto da avaliação das métricas e critérios.....	40
5.1.1	Extração das mensagens da sessão e cálculo das notas em cada métrica....	42
5.1.2	Geração do Relatório de Performance da Participação	46
5.1.3	Questionário de Avaliação do Relatório de Performance da Participação ...	47
5.2	Resultados e Conclusões sobre a Avaliação utilizando as Métricas.....	48

5.2.1	Quantidade Média de Caracteres por Mensagem – métrica mais aceita considerando a obtida nota	48
5.2.2	Tempo de Presença na Sessão – melhor entendimento e maior concordância com a métrica.....	51
5.2.3	Pertinência das Mensagens – métrica com menor aceitação	53
5.2.4	Análise da nota final – o Indicador da Participação	57
5.2.5	Entendimento das métricas individuais e da nota final (Indicador da Participação)	64
5.2.6	Concordância com as Métricas e Método para cálculo do Indicador da Participação.....	66
5.2.7	Concordância com as notas calculadas pela solução proposta	68
5.2.8	Comparando o Indicador da Participação com as notas atribuídas pelos próprios alunos na autoavaliação.....	70
5.2.9	Comparando a nota final com as notas atribuídas pelo Tutor e Autoavaliação	74
6.	Revisão das Métricas e Critérios.....	79
6.1	Problematizando as métricas propostas	79
6.2	Melhor definição de palavras chave para a sessão	80
6.3	Alteração da métrica Pertinência das Mensagens.....	82
6.4	Alteração do cálculo do Indicador da Participação	83
6.4.1	Aplicação de média aritmética para obtenção da nota final	85
6.4.2	Aplicação de média ponderada multiplicada pela presença no bate-papo ..	86
7.	Conclusão.....	89
7.1	Resultados Obtidos	89
7.2	Limitações e Trabalhos Futuros: desenvolvimento de sistema para análise da sessão e geração do Relatório de Performance da Participação	92
7.3	Limitações e Trabalhos Futuros: execução de outros estudos de caso para avaliar as métricas ajustadas	94
	Referências	96
	Apêndice A – Mapeamento de respostas dos alunos sobre a avaliação das métricas ..	100
	Apêndice B – Relatório de Performance da Participação.....	102

1. Introdução

O objetivo deste capítulo é descrever, de forma resumida, a motivação para pesquisa, além de apresentar uma visão geral da pesquisa e a organização do texto dessa dissertação. Na Seção 1.1 é justificada a presente pesquisa sobre o uso do bate-papo na educação a distância. A visão geral da pesquisa é apresentada na Seção 1.2 e, por fim, na Seção 1.3 é apresentada a organização do texto da dissertação, descrevendo a estrutura dos capítulos seguintes.

1.1 Bate-papo para Educação Online

A educação a distância tem crescido no Brasil e no Mundo, com aumento de 9% no número de matrículas no ano de 2011 em âmbito mundial (SANTOS *et al.*, 2012), e de acordo com dados do INEP/MEC (2012), já chegamos a mais de um milhão de graduandos brasileiros em cursos superiores nesta modalidade de ensino. Este número expressivo corresponde a mais de 15% do total de graduandos do Brasil (Figura 1).

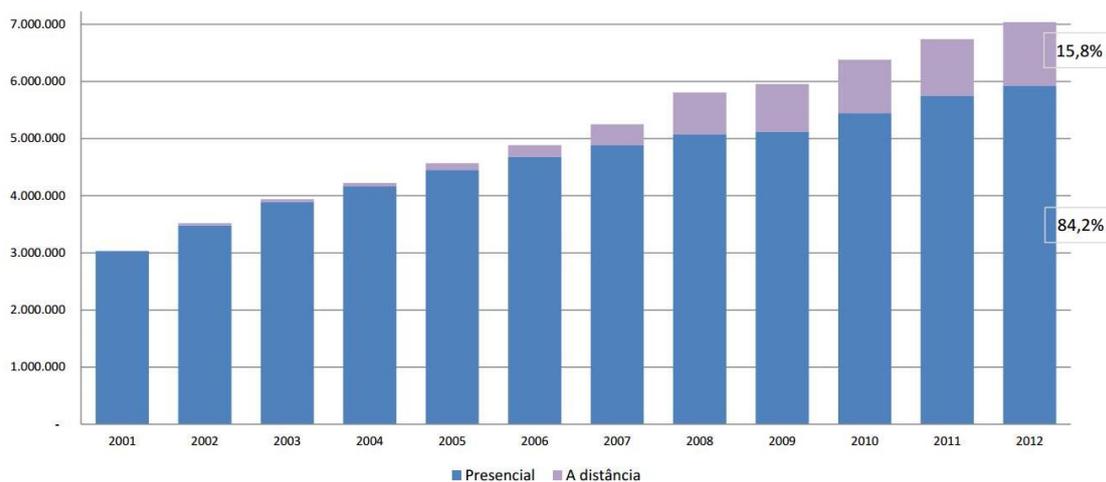


Figura 1. Evolução da Educação a Distância no Brasil

Dado o cenário do ensino a distância, é possível perceber que o serviço de bate-papo é um dos meios de comunicação mais utilizados para interação online, conforme apresentado na Figura 2, sendo adotado por mais de 60% das instituições brasileiras (CENSOEAD, 2012).

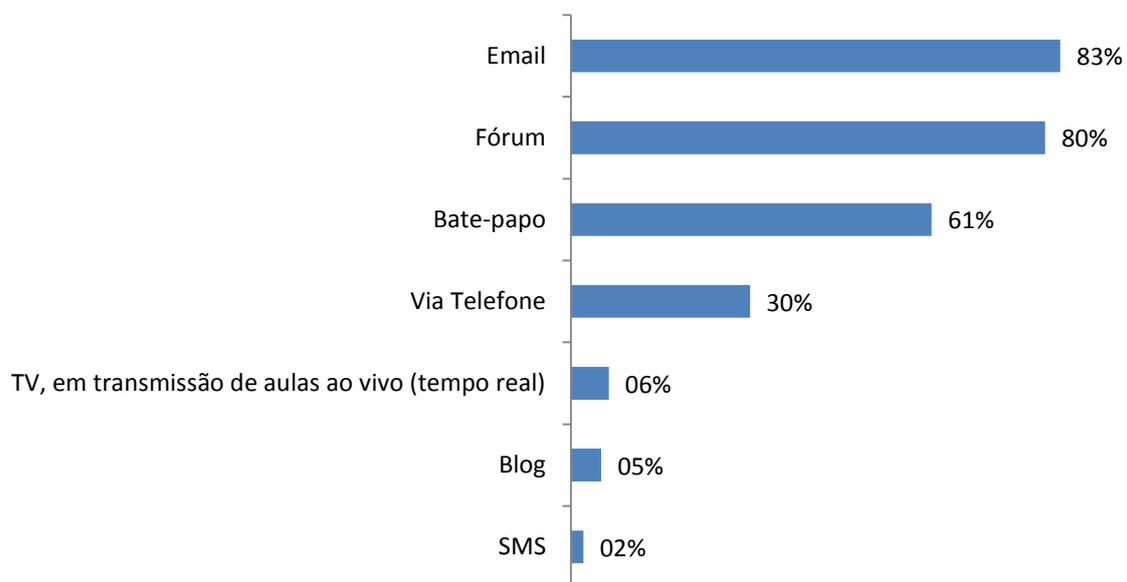


Figura 2. Meios de comunicação utilizados para interação online na EaD

O gráfico apresentado na Figura 2 mostra que os serviços de bate-papo têm considerável utilização na educação a distância. Porém, sua utilização tem ficado restrita para atendimento e tutoria, não aproveitando o potencial deste tipo de recurso. O bate-papo pode ser utilizado em aulas dinâmicas com possibilidade de discussão aberta entre os alunos e professor, onde cada um pode contribuir para a construção do conhecimento, expressar ideias e pontos de vista.

Em uma sessão de bate-papo normalmente os alunos podem expor seus pontos de vista de forma mais aberta do que percebe-se em salas de aula tradicionais, sem a visão do professor ocupando posição central, como detentor do conhecimento. A possibilidade de maior participação e redução da formalidade pode fomentar nos alunos o desejo de se expressar e dialogar, aumentando as chances do professor avaliar o conhecimento de cada aluno, dado que a avaliação depende do que é expressado como conhecimento pelo aluno.

O uso contínuo do bate-papo integrado, com atividades educacionais é também uma forma de manter a motivação dos alunos e o engajamento para garantir o sucesso e continuidade de um curso a distância (PIMENTEL, 2003a).

É sabido que a educação a distância pode ser aplicada mesmo que sem a utilização de recursos e sistemas online, com a utilização de material impresso, televisão, áudio etc. Considerando a adoção de serviços como E-mail, Fórum e o Bate-papo, percebe-se que a educação a distância está bem próxima da utilização plena de sistemas online. Por este motivo, iremos nos concentrar deste ponto em diante na educação online. Cabe, então, contextualizar o que se entende nesta pesquisa por educação online:

“A educação online é o conjunto de ações de ensino-aprendizagem ou atos de currículo mediados por interfaces digitais que potencializam práticas comunicacionais interativas e hipertextuais” (SANTOS, 2010, p. 5663).

Os Ambientes Virtuais de Aprendizado (AVA) são as tecnologias mais utilizadas nas práticas atuais de educação online, além das teleconferências e videoconferências (SANTOS, 2010). Alguns AVAs atuais possuem serviços para disponibilização de sessões de bate-papo educacional, o que reforça a utilização de bate-papo na educação online. Apesar da utilização de sessões de bate-papo no contexto da educação a online, os serviços que implementam o bate-papo, mesmo os disponíveis em Ambientes Virtuais de Aprendizagem (AVAs) como o Moodle, geralmente não possuem funcionalidades específicas para apoiar os educadores (ESTRUC *et al.*, 2012).

Assim como ocorre com outras atividades escolares, é importante avaliar os resultados obtidos com a aplicação de sessões de bate-papo. Uma das características a serem avaliadas na sessão de bate-papo é a participação dos alunos, dado que a participação pode estar relacionada à contribuição para a sessão.

Identifica-se que a avaliação da participação dos alunos nas sessões de bate-papo “é um instrumento adequado para equiparar as atividades realizadas no bate-papo com as outras atividades do curso” (FUKS *et al.*, 2006) e que a avaliação manual das sessões é uma atividade muito trabalhosa para os educadores (PIMENTEL *et al.*, Janeiro/2003), o que indica necessidade de apoiar esta atividade, visando auxiliar o trabalho dos educadores.

Dado este cenário, considerando o crescimento da adoção de sessões de bate-papo em meio educacional e o potencial que este tipo de serviço possui, o motivador desta pesquisa é a possibilidade de auxiliar os educadores na avaliação da participação dos alunos em sessões de bate-papo educacional.

1.2 Objetivo da pesquisa

O objetivo da pesquisa é apoiar os educadores na avaliação da participação de alunos em sessões de bate-papo educacional. Para apoiar a avaliação, nesta pesquisa buscou-se definir um conjunto de métricas objetivas e critérios para qualificar a participação dos alunos na sessão de bate-papo educacional, com a implementação de um Relatório de Performance da Participação, tendo como base as mensagens enviadas durante a sessão.

As **métricas** foram definidas para responder à pergunta: “O que considerar na avaliação?”. A definição das métricas permite identificar o que pode ser avaliado para classificar a participação dos alunos na sessão de bate-papo. Como exemplo de métricas, podemos citar “quantidade de mensagens enviadas durante a sessão” e “percentual de presença na sessão”. Todas as métricas definidas nesta pesquisa são objetivas, visto que se baseiam em valores reais, que podem ser extraídos automaticamente a partir do *log* da sessão de bate-papo, sem qualquer tipo de julgamento humano. Para cada métrica, os valores aferidos com base na sessão é o que chamamos de **medidas**. Por exemplo, as 10 mensagens enviadas durante a sessão, por determinado participante, é o que chamaremos de **medida** da métrica “quantidade de mensagens enviadas durante a sessão”.

Com a extração das métricas da sessão de bate-papo temos a apresentação da realidade sobre a sessão e torna-se necessário qualificar os dados levantados através da comparação com os resultados desejados – ponto onde introduzimos os **critérios**. Os critérios correspondem às expectativas que existem para cada métrica definida no processo de qualificação da participação e servirão como base para comparação com a realidade vista na sessão de bate-papo, no caso desta pesquisa, permitindo a identificação do nível de participação dos alunos.

Por exemplo, poderíamos considerar que o **critério** para avaliação da **métrica** “quantidade de mensagens enviadas durante a sessão” seria qualificar com nota máxima

de participação os alunos que enviassem mais que 30 mensagens (que seria a **medida** aferida do aluno para respectiva métrica).

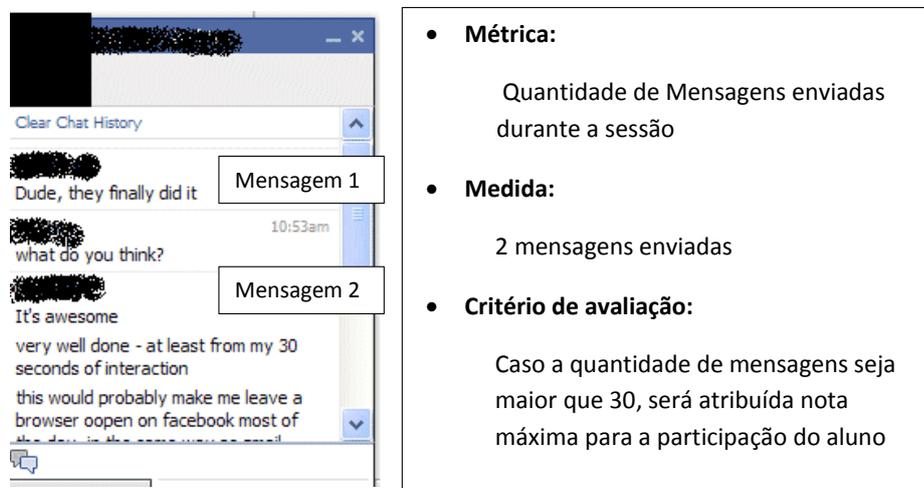


Figura 3. Exemplo de Métrica, Medida e Critério

O Relatório de Performance da Participação criado nesta pesquisa reúne as métricas, medidas e critérios, introduzidos anteriormente, qualificando a participação dos alunos na sessão de bate-papo em questão. Tal qualificação (consolidada com o Relatório de Performance da Participação) é o que servirá de apoio aos educadores para a avaliação da participação dos alunos na sessão de bate-papo.

1.3 Visão Geral da Pesquisa

1.3.1 Apresentação do Problema

A análise manual de sessões de bate-papo é difícil, considerando a quantidade de mensagens enviadas e o tempo necessário para a análise manual de tais mensagens. Para que o avaliador possa definir as notas dos participantes seria necessário analisar cada mensagem para identificação da contribuição dos alunos. Os sistemas atuais não oferecem apoio à análise e em alguns casos é trabalhoso inclusive recuperar as mensagens enviadas.

1.3.2 Proposta de Solução

O primeiro passo da proposta de solução foi o planejamento, com base na literatura, das métricas objetivas que seriam utilizadas para a qualificação da participação dos alunos na sessão de bate-papo. Foram selecionadas quatro métricas

para qualificação da participação: “Quantidade de Mensagens”, “Quantidade de Caracteres por Mensagem”, “Pertinência das Mensagens” e “Tempo de presença dos Alunos na Sessão”.

Para cada métrica foram extraídas as medidas de cada aluno e considerados critérios de avaliação para qualificar os alunos em cada métrica, utilizando um percentual de aproveitamento. O resultado final esperado é a geração de um Indicador da Participação no bate-papo para cada aluno, calculado através da multiplicação do percentual de cada métrica individual aferida.

Tendo como base as métricas selecionadas foi criado o relatório de Performance da Participação, que apesar de ter sido gerado manualmente nesta pesquisa poderia também ser gerado de forma automática através a implementação de um sistema para tal objetivo, utilizando as métricas e os critérios definidos, além das medidas aferidas durante a sessão de bate-papo.

1.3.3 Avaliação da Solução

Foi conduzido um estudo de caso exploratório para avaliação da solução proposta, entendimento da aplicabilidade e aceitação das métricas, critérios e do Relatório de Performance da Participação propostos. Neste estudo de caso foi conduzida uma sessão de bate-papo com aproximadamente 1 hora de duração e 15 participantes, além do tutor.

Após a sessão realizada, os dados foram extraídos e analisados. Foram calculados indicadores individuais para geração das notas dos participantes, como quantidade de mensagens e caracteres enviados por cada participante, horário da primeira mensagem enviada por cada aluno e o tempo de permanência na sessão. Estas informações serviram de base para a geração da nota nas métricas “Quantidade de Mensagens”, “Quantidade de Caracteres por Mensagem” e “Presença na Sessão”.

Para a avaliação da métrica “Pertinência das Mensagens” foi utilizada uma estratégia para tratamento das mensagens enviadas e a identificação de palavras-chave. Os dados gerados foram utilizados como entrada para uma rotina de programação criada para geração do percentual de mensagens pertinentes enviadas.

Com a recuperação de todas as notas individuais dos participantes em todas métricas definidas foi possível aplicar a fórmula para geração da nota final e geração do Relatório de Performance da Participação.

A avaliação da solução foi realizada através de questionário sobre o formato de avaliação, com perguntas relativas ao entendimento das métricas e critérios utilizados e concordância com as notas recebidas, além de questões livres para comentários gerais. Os dados do questionário foram avaliados para entendimento da aceitação do método de qualificação da participação proposto nesta pesquisa, métricas e notas atribuídas aos alunos, além de oportunidades de melhorias e ajustes necessários.

1.4 Organização do texto da Dissertação

O capítulo 2 apresenta uma discussão sobre o processo de avaliação da participação em sessões de bate-papo, analisando as conexões entre a avaliação escolar, a participação em bate-papo e a avaliação da participação em sessões de bate-papo.

No capítulo 3 são apresentados os problemas e dificuldades relacionadas à avaliação da participação em sessões de bate-papo, considerando especialmente as sessões voltadas para educação.

Após introduzir os principais conceitos nos capítulos 2 e 3, o capítulo 4 apresenta a proposta de solução, detalhando as métricas selecionadas e os critérios adotados para cálculo das notas e geração do Relatório de Performance da Participação.

O capítulo 5 apresenta em detalhes o estudo de caso realizado para avaliar as métricas definidas, explicadas no capítulo 4. Ainda no capítulo 5 são abordados alguns resultados e conclusões alcançados com o estudo de caso.

Após o estudo de caso foi possível perceber oportunidades de melhoria e necessidades de ajustes nas métricas e critérios de avaliação. O capítulo 6 apresenta a revisão das métricas e os ajustes que poderiam ser realizados para novos estudos.

Por fim, o capítulo 7 apresenta as conclusões do trabalho, bem como os resultados obtidos, indicações de limitações enfrentadas e possíveis trabalhos futuros, para continuidade da pesquisa.

2. Avaliação da Participação em Bate-papo

O objetivo deste capítulo é discutir sobre o processo de avaliação dos alunos em sessões de bate-papo. Para tanto, é necessário entender o que se considera como avaliação e seus objetivos. Este capítulo aborda o tema referente à Avaliação da Participação em sessões de Bate-papo (seção 2.4), passando primeiramente por uma discussão teórica sobre os temas Avaliação na Educação (seção 2.1), Avaliação na Educação On Line (seção 2.2) e Avaliação da Participação na Educação On Line (seção 2.3).

2.1 Avaliação na Educação

A avaliação da aprendizagem é um importante aspecto para a evolução dos alunos e para o entendimento do progresso individual. A avaliação é parte do ato pedagógico, que por sua vez é composto por Planejamento, Execução e Avaliação, conforme explica LUCKESI (2011):

O ato pedagógico – composto de três elementos – inicia-se com o estabelecimento de metas, com o planejamento. Sob sua guia, segue a execução, que dialeticamente, soma-se à avaliação, para que se produza o resultado desejado. (LUCKESI, 2011, p. 19)

É importante planejar e definir metas para que seja possível avaliar o alcance de tais metas. Em outras palavras, podemos dizer que primeiro precisamos identificar o que queremos alcançar com a ação pedagógica para então executar as ações e avaliar os resultados alcançados. Sem a avaliação, baseada no planejamento e definição de metas, os resultados alcançados serão aleatórios e, possivelmente, insatisfatórios (LUCKESI, 2011).

A prática da avaliação precisa ser bem pensada e assertiva, pois possui considerável poder no comportamento e evolução dos educandos, visto que pode ser encarada como um feedback pelos alunos. Avaliar negativamente sem a apresentação de pontos efetivos de melhorias, de forma construtiva, pode reduzir o empenho e motivação, causando um resultado negativo. Por outro lado, caso a avaliação seja bem

sucedida no objetivo de apontar possibilidades de melhoria e seja compreendida pelo aluno, teremos a oportunidade de impactá-lo positivamente.

O processo de fornecer feedback sobre a aprendizagem está no cerne do que chamamos de Avaliação Formativa. De acordo com (PERRENOUD, 1999) podemos considerar como formativa “toda prática de avaliação contínua que pretenda contribuir para melhorar as aprendizagens em curso, qualquer que seja o quadro e qualquer que seja a extensão concreta da diferenciação do ensino”. Este tipo de avaliação tem como característica a interatividade, centrada nos processos cognitivos dos alunos e associada aos processos de feedback, de regulação, de autoavaliação e de auto-regulação das aprendizagens (FERNANDES, 2006).

Neste processo de entendimento da aprendizagem, cabe ao educador fornecer feedback sobre a aprendizagem dos alunos para orientação dos pontos de melhoria individuais. O entendimento do progresso é um ponto importante para que os alunos possam investir em sua aprendizagem individual. Do ponto de vista do educador, a avaliação serve com recurso para apoiar o aluno no seu crescimento e aprendizagem, além de permitir ao educador avaliar seu método de ensino, permitindo correções de rumos da atividade e de seus resultados (LUCKESI, 2011).

Outra fonte de regulação da aprendizagem, além da avaliação comum, é o processo de autoavaliação, que corresponde a um julgamento pessoal do próprio indivíduo. Na autoavaliação o próprio aluno toma consciência dos diferentes momentos e aspectos da sua atividade cognitiva (SANTOS, 2002). O olhar interno para suas próprias respostas para os temas estudados pode ajudar os educandos a entender seus pontos fracos e necessidades de melhorias, motivando-os a aprofundar nos assuntos que precisa melhorar seus conhecimentos.

A prática de autoavaliação pode ser combinada com as práticas de avaliação, visando melhorar a assertividade do processo de avaliação. Em PERRENOUD & THURLER (2002), sugere-se a combinação dos dois tipos de condutas:

- A autoavaliação tem como objetivo identificar o impacto das ações empreendidas a fim de determinar a sequência das operações, bem como os ajustes eventualmente necessários;
- A avaliação externa verifica a pertinência da autoavaliação e utiliza os resultados desta para definir as linhas gerais da ação do sistema.

Para o educador, utilizar métodos adequados para avaliar a aprendizagem dos alunos é crucial no caminho de garantir os resultados desejados. Para os alunos, avaliações coerentes e construtivas permitem o melhor conhecimento de seus pontos fracos e fortes, proporcionando uma oportunidade de melhorar e monitorar seu avanço e conhecimento. Do ponto de vista da educação, de forma geral, a avaliação permite que o processo seja alimentado com experiências positivas e negativas, que irão auxiliar na busca pelo objetivo de formar melhores indivíduos.

2.2 Avaliação na Educação online

A evolução da educação traz mudanças no dia-a-dia dos educadores, promovendo novas visões e opções para atacar os problemas relacionados ao ensino. As práticas de avaliação precisam seguir os avanços para manterem-se atualizadas e coerentes com as novidades que se apresentam. Os educadores precisam estar atentos às mudanças para acompanhar a evolução aplicarem avaliações que alcancem o objetivo de informar as deficiências e os pontos positivos para contribuir para a formação dos alunos.

O advento da Educação a Distância apresenta uma nova perspectiva de avaliação, que requer rupturas com o modelo tradicional de avaliação (SILVA, 2006). É necessário repensar as técnicas de avaliação considerando a nova perspectiva, de forma que os objetivos da avaliação sejam alcançados, conforme discutido por CALDEIRA (2004).

Os ambientes digitais de aprendizagem possuem elementos que se configuram como um novo contexto educacional, diverso do presencial, e por isso é fundamental que se criem processos e estratégias que respondam às novas necessidades e circunstâncias dos novos modelos. Não é possível, portanto, simplesmente adaptar os modelos presenciais. (CALDEIRA, 2004, p. 6)

Cada vez mais torna-se importante a participação dos alunos e professores no processo de aprendizagem, de forma que as experiências contribuam para o aprendizado e que o cenário onde o aluno apenas é um receptor de informações seja minimizado, prevalecendo a participação e a co-autoria. Com esta abordagem espera-se que o aluno tenha mais participação no processo de aprendizagem, onde professor e aluno formem uma rede e não uma rota, de acordo com SILVA (2006).

Professor e alunos constroem uma rede e não uma rota. Ele define um conjunto de territórios a explorar. E a aprendizagem e a avaliação se dão na exploração – ter a experiência de participar, de colaborar, de criar, de co-

criação – realizada pelos aprendizes e não a partir da sua récita, do seu falar-ditar. Isso significa modificação no clássico posicionamento na sala de aula. (SILVA, 2006 p. 32)

A avaliação dos alunos na educação online não é uma tarefa fácil para os professores, dado que a maioria dos ambientes virtuais utilizados não oferecem recursos para apoio na avaliação (DA ROCHA, 2006), restringindo-se ao registo das informações trocadas. Nestes casos, é necessário realizar as avaliações de forma manual analisando dados quantitativos extraídos das ferramentas com a análise manual das interações dos alunos, o que gera trabalho e consome tempo dos educadores.

Para auxiliar na avaliação, alguns estudos sugerem a aplicação de técnicas de mineração de texto e dados, conforme foi proposto em (SILVA *et al*, 2001), onde é sugerido o acompanhamento das atividades durante a utilização do ambiente virtual da turma em níveis:

- 1- Rastreamento de páginas visitadas e mensagens trocadas;
- 2- coleta dos resultados das atividades propostas e
- 3- análise dos dados coletados nos níveis anteriores. A análise dos dados coletados adota algoritmos de mineração de dados para buscar padrões de comportamento dos aprendizes (por exemplo, encontrar comportamentos que caracterizam aprendizes aprovados). (DA ROCHA, 2006).

A utilização de técnicas de mineração de dados também é explorada em (ZAIANE & LUO, 2001), onde foi criado um framework para utilização de técnicas de mineração de dados para recuperar e avaliar *logs* de acesso do ambiente virtual, baseado em três passos: coleta de dados e pré-processamento (para organizar os dados que serão avaliados), descoberta de padrões (utilização de algoritmos para descoberta de padrões relevantes e potencialmente úteis) e análise dos padrões (onde ocorre a análise dos padrões descobertos e momento onde ocorrem as conclusões). No caso da realização de sessões de bate-papo, a matéria prima a ser analisada é a discussão realizada durante a sessão, principalmente as mensagens enviadas pelos alunos durante a sessão.

O ambiente TelEduc é um sistema para realização cursos a distância oferecidos via Internet, desenvolvido pelo NIED (Núcleo de Informática Aplicada a Educação) da Universidade Estadual de Campinas (Unicamp), que disponibiliza funcionalidades para apoio aos educadores na avaliação dos alunos, como o Intermap e Acessos. O Intermap permite o mapeamento da interação e participação dos alunos através de gráficos

utilizando técnicas de visualização de informação, enquanto o Acessos permite a geração de relatórios com a possibilidade de extração de informações como número de acessos, data e hora, frequência e o acesso a cada uma das ferramentas do TelEduc, como chat e fórum (DA ROCHA, 2006).

O TelEduc apresenta ainda relatórios de acompanhamento para apoio na análise e avaliação formativa dos alunos: Relatório de Participação (permite a identificação de alunos ativos e passivos na atividade), Relatório de Verificação de Notas e Relatório de Notas da Turma (que permite a verificação do desempenho dos aprendizes e a visão geral da turma) e Relatório de Histórico do Desempenho (que permite a visualização dos aprendizes nas atividades, sendo um importante artefato para a avaliação formativa) (DA ROCHA, 2006). As funcionalidades oferecidas pelo TelEduc representam um considerável apoio para a avaliação dos alunos na educação online, oferecendo aos educadores e alunos uma oportunidade de melhorar o processo de ensino.

O ambiente AulaNet, um ambiente para a criação, participação e administração de cursos baseados na Web (FUKS, 2000), também foi experimentado para a educação online, com serviços para a composição de curso, que parte do princípio que para aprender em grupo, um indivíduo tem de se comunicar, se coordenar e cooperar (FUKS et al, 2006), baseado no Modelo 3C de comunicação (ELLIS et al, 1991). O ambiente foi experimentado com turmas reais totalmente a distância.

No AulaNet também foram disponibilizadas opções para a avaliação dos alunos, como funcionalidade para a criação de exames para avaliação dos aprendizes pelos educadores. O ambiente disponibilizava também um relatório da participação dos alunos nos diversos eventos dos cursos aplicados, como em grupos de discussão, por exemplo (FUKS, 2000).

A avaliação da aprendizagem dos alunos em ambientes de educação online considera a avaliação da participação durante o curso. Não basta avaliar os alunos pela quantidade de participações, mas torna-se necessário também avaliar a qualidade da participação, para que seja possível identificar o grau de envolvimento com o tema e o nível de aprendizagem dos alunos. Na próxima seção, será discutido com mais detalhes o processo de avaliação da participação na educação online.

2.3 Avaliação da Participação na Educação Online

A participação dos alunos no processo de aprendizagem é importante para qualquer paradigma de ensino. Parte do sucesso do processo de aprendizagem em meio digital está diretamente ligado à participação e integração social da comunidade de aprendizagem (BARILLI, 2006). Em sessões de bate-papo, por exemplo, podemos perceber que esta participação é essencial, visto que possibilita a troca de experiências sobre os temas discutidos. Assim, avaliar a participação parece uma necessidade para entender o grau de contribuição dos alunos para as discussões.

Sobre os aspectos que devem compor a avaliação dos alunos devemos considerar a inclusão de critérios quantitativos e qualitativos, que não são excludentes, mas sim complementares e a quantificação só se complementa com a avaliação qualitativa (NUNES e VILARINHO, 2006). Indicadores quantitativos podem ser trabalhados a partir de eventos ocorridos durante a sessão, como a quantidade de mensagens enviadas, enquanto indicadores qualitativos podem ser extraídos a partir da relevância das mensagens enviadas para o tema em discussão.

Em (BARILLI, 2006) foi utilizado um modelo baseado em critérios para avaliação da participação em Fóruns. Os critérios utilizados foram: Consistência das Colocações (critério qualitativo), Clareza e concisão (critério qualitativo) e Interação (critério quantitativo, considerando a quantidade de participações), onde para cada critério foi atribuída uma quantidade de pontos (de 0 a 10, de acordo com o aproveitamento do aluno). Foram estabelecidos conceitos para avaliação da pontuação dos alunos que variaram de “A” (entre 25 e 30 pontos) a “D” (entre 0 e 14 pontos). Neste modelo, foram mesclados critérios qualitativos e quantitativos para atender aos objetivos da avaliação.

Também podemos qualificar a participação dos alunos em sessões de bate-papo para identificar, por exemplo, a formação de grupos de discussão e avaliar os alunos mais engajados com a sessão. Algumas abordagens para avaliação da participação em sessões de bate-papo serão apresentadas na seção a seguir.

2.4 Avaliação da Participação em sessões de Bate-Papo educacional

Foi apresentado neste trabalho que a avaliação dos alunos é de grande importância para o aprendizado na educação online ou tradicional. Avaliar a participação dos alunos

nas sessões de bate-papo e a avaliação das mensagens enviadas é também importante, pois incentiva a participação dos alunos e o aumento da qualidade das mensagens (PIMENTEL et al, 2004).

Durante a sessão de bate-papo os alunos podem participar de formas diferentes, com mais ou menos envolvimento na discussão e contribuição para o tema. Em (HRASTINSKI, 2008) a participação foi classificada em seis diferentes níveis, de acordo com o tipo de envolvimento dos presentes:

- Nível 1: Acesso ao ambiente virtual
- Nível 2: Participação escrita
- Nível 3: Participação com qualidade da escrita
- Nível 4: Participação com escrita e leitura
- Nível 5: Participação com escrita real e percebida
- Nível 6: Participação como se juntando a um diálogo

Podemos considerar que ocorre participação nos seis níveis indicados pelo autor, com possibilidade de avaliação dos alunos que se enquadram em cada nível. Um aluno que esteja enquadrado no primeiro nível de participação deveria receber avaliação com menor pontuação que um aluno que esteja no sexto nível. Assim, a avaliação dos educadores baseado nestes níveis poderia atribuir uma nota de aproveitamento para cada nível de participação.

A classificação dos alunos nos níveis de participação pode ser feita analisando a própria discussão. Um aluno que entrou na sala de bate-papo, mas não enviou mensagens pode ser enquadrado no nível 1, enquanto um aluno com ao menos uma mensagem enviada poderia ser enquadrado no nível 2. Apenas entrar na sessão ou enviar mensagens de socialização não seria o ideal para uma sessão, pois este comportamento não necessariamente contribui para a qualidade da sessão, que começa a ser alcançada com participação no nível 3, onde as mensagens enviadas contribuem para a discussão. A evolução da participação continua ocorrendo até o nível 6, onde o aluno consegue acompanhar toda discussão e se sente co-responsável pela produção de mensagens de qualidade.

A evolução da participação na sessão pode ser influenciada por características da sessão, como a quantidade de participantes, por exemplo, que pode interferir na

qualidade da conversa. Em uma sessão de bate-papo educacional é importante que todos consigam ler todas as mensagens e que também possam enviar mensagens, motivo pelo qual é preciso limitar a quantidade de participantes (ROCHA, 2013). O trabalho apresentado em (ROCHA, 2013) atacou justamente este ponto, com o objetivo de estimar a quantidade de alunos que devem participar de uma sessão de bate-papo para que seja alcançado um nível satisfatório de participação, atingindo o nível 6, de preferência.

A abordagem utilizada para avaliação da participação em sessões de bate-papo apresentada em (FUKS *et al*, 2006) foi a ponderação da qualidade pela quantidade de mensagens em cada sessão de bate-papo realizada no ambiente AulaNet. O processo consiste na avaliação da qualidade das mensagens pelo moderador da sessão (através da análise humana do texto, com a atribuição de nota) e ponderação da média de notas do aluno na sessão pela quantidade de mensagens enviadas pelo aluno na mesma sessão. A ponderação e análise qualitativa visa equilibrar os dois aspectos, evitando que um aluno que enviou muitas mensagens, sem qualidade, obtenha nota máxima, assim como um aluno que enviou poucas mensagens, com qualidade, também não alcance um patamar máximo na nota de participação (visto que participou pouco da discussão). A explicação para esta abordagem de avaliação da participação pode ser apresentada com o seguinte trecho:

Numa boa discussão, todos devem dar contribuições significativas e de forma equiparável. Quem participa sem se preparar ilude e desaponta o grupo. Quem apenas lê, escuta ou vê as mensagens de uma conferência não está participando, apenas assistindo. Quando uma pessoa participa muito mais do que as outras, está monopolizando a discussão. (FUKS et al, 2006)

Algumas opções definidas em FUKS et al (2006) para a ponderação da avaliação da participação em função da quantidade estão ilustradas na Figura 4.



a) Quantidade irrelevante
A quantidade de mensagens não influencia na nota (avaliação apenas qualitativa)

b) Quantidade mínima
O peso aumenta com a quantidade de mensagens até atingir um máximo

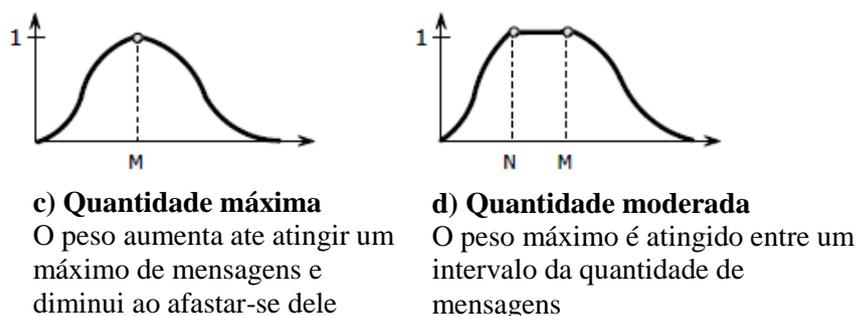


Figura 4. Opções para avaliação da participação

A opção “a” desconsidera a quantidade de mensagens e avalia a participação apenas de forma qualitativa. A opção “b” atribui a nota de acordo com o alcance de um volume de mensagens, mesmo que a quantidade de mensagens seja muito alta (caso em que o participante poderia estar monopolizando a discussão e não seria penalizado por isso). Na opção “c” a nota máxima é atribuída apenas para alunos que alcançaram o limite de mensagens e varia de acordo com a aproximação ou afastamento da quantidade de mensagens desejada. Por fim, a opção “d” permite atribuir a nota máxima para alunos que atingiram um determinado intervalo de quantidade de mensagens e parece ser o mais adequado, visto que define um intervalo para a quantidade desejável de mensagens e reduz a nota gradativamente caso o aluno esteja abaixo ou acima do limite (penalizando alunos com excesso de participação ou com baixa participação).

Algumas técnicas podem ser utilizadas para avaliar a participação em discussões on-line, mesmo fora do ambiente educacional, como a aplicação de técnicas de análise de redes sociais e utilização de medidas simples disponíveis em sistemas de análise de redes sociais, como a UCINET e NetMiner. O trabalho apresentado em (RABBANY *et al.*, 2011) introduziu sua ferramenta para avaliação da participação em discussões on-line: o Meerkat-ED. O objetivo do sistema é gerar visualizações para a participação dos envolvidos, suas interações e os líderes da discussão, bem como alunos periféricos (RABBANY *et al.*, 2011).

A visualização de redes sociais na discussão pode auxiliar na avaliação da participação apresentando a proximidade dos alunos com os tópicos em discussão e frases enviadas, além da proximidade dos alunos, identificando os grupos mais interessados na discussão.

A aplicação de métodos de mineração de dados também pode apoiar na avaliação de sessões de bate-papo. A classificação das mensagens pode ser realizada através de métodos de mineração de dados de forma automática e em tempo real, oferecendo possibilidades de melhorias nos ambientes utilizados para aprendizado (ANJEWIERDEN *et al.*, 2007). A classificação visual de mensagens por assunto pode auxiliar o moderador da sessão a identificar se o tópico está sendo abordado ou se está ocorrendo uma dispersão na discussão.

Abordagens mais próximas de mineração de texto também estão em utilização para a análise de bate-papo. O sistema proposto em (HUI *et al.*, 2008) permite a análise inteligente de mensagens de bate-papo utilizando técnicas de mineração de texto, com funcionalidades para identificação de redes sociais no bate-papo e interações dos usuários, além de detecção automática do tópico da sessão.

A extração de informações de sessões de bate-papo são também abordadas em (GAINARU *et al.*, 2010), que apresenta um sistema para tal fim. O sistema utiliza processamento de linguagem natural para análise automática de sessões de bate-papo, permitindo a detecção dos tópicos e relacionamentos entre as mensagens. Um dos objetivos é permitir a identificação do grau de comprometimento dos participantes com a discussão.

No estudo apresentado em (PIMENTEL *et al.*, 2003b) foram aplicados 3 variações de métodos para avaliação da participação em sessões de bate-papo, a saber:

- **Método 1 - Conceito em função da quantidade relativa de mensagens:** realização da avaliação baseada na quantidade de mensagens enviadas pelos participantes durante a sessão, classificando os alunos nos conceitos: Muito Ativo, Ativo, Pouco Ativo e Desinteressado.

A atribuição do conceito depende de análise humana para identificação da quantidade de mensagens que enquadraria os alunos em cada conceito, visto que pode variar para cada sessão. A avaliação da qualidade das mensagens (remoção de mensagens fora de contexto, por exemplo) também precisa ser manual, o que dificulta a utilização do método.

- **Método 2 - Nota em função da quantidade absoluta de mensagens:** método que calcula a nota do aluno através da quantidade absoluta de mensagens

enviadas durante a sessão. A quantidade normal de mensagens foi calculada considerando estudos anteriores e a classificação dos alunos é realizada de acordo com a quantidade de mensagens enviadas (se for superior à quantidade normal é atribuída nota máxima e, caso contrário, é atribuída uma nota proporcional de acordo com a aproximação da quantidade normal de mensagens).

Diferente do método 1, o método 2 pode ser automatizado para gerar as notas com apoio computacional, sem análise humana como ocorre com o método 1.

- **Método 3 - Nota em função da qualidade ponderada pela quantidade:** neste método todas as mensagens são analisadas para classificação de conceitos para cada mensagem (Bom – nota 10, Regular – nota 5 e Péssimo – nota 0) de acordo com a pertinência com o tema em discussão. A nota foi calculada com a média das notas das mensagens ponderada por um peso calculado em função da quantidade de mensagens (o peso foi calculado utilizando o mesmo critério do método 2).

Este método foi considerado o mais adequado, pois apesar de mais trabalhoso gerou notas mais adequadas na opinião dos pesquisadores e dos alunos que participaram das sessões avaliadas por este método.

A presente pesquisa considerou também um método para a avaliação da participação dos alunos na sessão de bate-papo, com base em um conjunto de métricas que devem ser observadas, juntamente com critérios, para atribuição das notas dos alunos, visando gerar um indicador da participação de cada aluno em uma sessão de bate-papo. O método proposto e os critérios considerados nesta pesquisa se baseiam no conhecimento gerado pelos trabalhos anteriores, apresentados neste capítulo, além de suposições feitas durante o estudo da literatura.

3. Problematizando a Avaliação da Participação em Bate-Papo

A avaliação é um processo importante para a garantia da aprendizagem dos alunos e exige conhecimento, por parte dos professores, sobre os resultados das ações pedagógicas. Em turmas tradicionais os educadores utilizam como meios para entender a realidade dos alunos artefatos como exames, testes, etc. Em sessões de bate-papo, a principal informação a ser utilizada são as mensagens enviadas pelos alunos durante a sessão. Existem algumas dificuldades para a análise destas mensagens (que tem por objetivo gerar informações e indicadores sobre a sessão e sobre os alunos), que serão abordadas nas próximas seções deste capítulo.

3.1. Dificuldades na avaliação da participação em sessões de bate-papo

As sessões de bate-papo correspondem a um recurso que permite a aprendizagem colaborativa. Na aprendizagem colaborativa há a valorização da participação dos alunos e sua capacidade de resolução de problemas (FUKS *et al.*, 2006). Caso sejam utilizadas sessões de bate-papo como recurso para aprendizagem colaborativa e não seja possível avaliar a participação dos alunos não teremos formas de medir se o objetivo da sessão foi alcançado.

Os sistemas atuais não oferecem mecanismos para a avaliação da participação dos alunos nas sessões de bate-papo, bem como a estrutura das interações desenvolvidas. Em alguns casos apresentam informações estatísticas não úteis para a avaliação da participação, obrigando os educadores a analisar manualmente as sessões, de forma trabalhosa e sujeita a erros (RABBANY *et al.*, 2011).

O trabalho necessário para a avaliação manual das sessões de bate-papo, de modo geral ou apenas da participação, pode causar subutilização e até abandono deste tipo de sistema pelos educadores. Considerando a utilização de sessões de bate-papo na educação online, conforme discutido nos capítulos anteriores, podemos ter perda considerável caso não tenhamos sistemas de apoio que auxiliem os educadores na utilização deste tipo de recurso.

Os tópicos a seguir tratam de alguns problemas típicos enfrentados na avaliação da participação em sessões de bate-papo.

3.2. A análise humana de mensagens enviadas durante a sessão de bate-papo consome muito tempo dos educadores

Após a realização de uma sessão de bate-papo, os educadores responsáveis precisarão avaliar os alunos envolvidos. Foram apresentados nesta pesquisa trabalhos onde tal avaliação é realizada de forma manual, com análise humana de cada mensagem. Logo, para a avaliação das mensagens o educador deveria recuperar as mensagens para leitura ou acessar o sistema de bate-papo para reler as mensagens enviadas. Caso o educador não tenha participado do bate-papo, será mais difícil avaliar a participação dos alunos visto que não terá o contexto da discussão e o sentimento de como foi conduzida.

Alguns AVAs permitem o uso de serviços como Fóruns, Questionários e Sala de Bate-Papo para o trabalho com os temas compreendidos nas ementas. O Moodle, por exemplo, permite a inclusão de módulo específico para Bate-Papo, com funcionalidade administrativa para exportação das sessões, que facilita a extração dos dados pelos moderadores, mas não fornece recursos específicos para apoio a avaliação dos alunos na sessão.

A necessidade de tratamento das mensagens da sessão, para arquivamento e utilização posterior ou para análise da participação, se apresenta como mais um ponto de dificuldade para os educadores.

O estudo apresentado em (ROCHA, 2013) exhibe um resumo de dados coletados em 142 sessões de bate-papo em meio educacional do curso TIAE (Tecnologias de Informação Aplicadas à Educação, ministrada no Departamento de Informática da PUC-Rio). Considerando este estudo, nas 142 sessões, foram enviadas em média 320 mensagens por sessão (com média de 35 mensagens por participante) e 72 caracteres por mensagem.

Dado o volume de informações médio nas sessões, caso o moderador necessite de 30 segundos para analisar cada mensagem, seriam necessárias mais de 2 horas para análise de uma sessão de 320 mensagens. Com o tempo de sessão médio de 1 hora

indica que o moderador gastaria mais de duas vezes o tempo da sessão para analisar as mensagens, sem considerar o trabalho necessário para gerar as avaliações individuais.

A análise automática ou semiautomática das sessões poderia auxiliar na geração de estatísticas individuais sobre cada participante, permitindo que o moderador analisasse os dados, sem necessidade da análise de mensagem por mensagem.

3.3. Dificuldades para a análise automática (ou semiautomática) das mensagens

Dada a dificuldade para analisar manualmente as mensagens das sessões de bate-papo, algumas técnicas estão sendo aplicadas para a realização de análises automáticas das sessões para entendimento do tópico em discussão, vocabulário utilizado e outros exemplos. Além da quantidade de texto produzido, que dificulta a análise, alguns problemas específicos das sessões de bate-papo podem dificultar também a análise automática das conversas.

As mensagens em bate-papo são normalmente curtas e concisas, com menos de 10 palavras, em geral, e tal característica causa dificuldades na detecção do tópico da discussão (HUI *et al.*, 2008). Para minimizar os problemas, em alguns casos, são usados recursos para entendimento do contexto de palavras utilizadas e relacionamentos semânticos (ADAMS & MARTELL, 2008), visando verificar se a discussão está adequada ao tema principal desejado.

A análise dos dados em salas de bate-papo enfrenta também problemas característicos deste tipo de sistemas, como o uso de gírias, abreviaturas e a ocorrência de erros ortográficos. Tais características dificultam a execução de métodos para análise das mensagens e causam distorção nos resultados (LOH *et al.*, 2004).

Em alguns casos são combinadas técnicas de mineração e processamento de texto para redução dos problemas. Um pré-processamento com o uso de técnicas como *tokenização*, eliminação de *stop words*, *stemming* permite gerar uma nova massa de dados para análise, com melhor acurácia no resultado final (GAINARU *et al.*, 2010).

Um dos desafios da avaliação automática ou semiautomática é a descoberta da qualidade das mensagens. A análise quantitativa baseia-se na identificação da quantidade de mensagens e caracteres enviados, por exemplo, para aferir a quantidade

de texto produzido. Como não necessariamente enviar muitas mensagens é sinônimo de participação e de boa participação, é necessário avaliar a qualidade das mensagens.

A avaliação qualitativa de forma semiautomática ou automática das mensagens enviadas depende da análise do tema em discussão. A descoberta do tema da sessão e a pertinência das mensagens enviadas em função do tema poderiam gerar a pontuação para os alunos baseada na quantidade e na qualidade das mensagens, mais próxima do modelo ideal para a avaliação da participação.

4. Proposta de Solução: Métricas Objetivas e Critérios para apoiar a Avaliação da Participação em sessão de bate-papo educacional

O ato de avaliar pode ser explicado por dois processos articulados e indissociáveis: “Diagnosticar” e, se necessário, “Intervir, tendo em vista a melhoria dos resultados”. No primeiro momento, o Diagnóstico, permite qualificar a realidade por meio de sua descrição considerando dados relevantes e critérios para comparação com a qualidade desejada. No segundo passo, a Intervenção na Realidade, há a tomada de posição, conduzindo à intervenção, se necessária, para corrigir o rumo para que os objetivos e metas definidos sejam alcançados (LUKCESI, 2011). Neste ponto, fica claro que o Diagnóstico correto apoia a Intervenção e correção de rumo, quando necessária.

Para apoiar a avaliação do professor sobre a participação dos alunos numa sessão de bate-papo, nesta pesquisa foram propostas métricas objetivas para diagnosticar a participação dos alunos (Descrição da Realidade) e critérios para qualificação desta participação (Qualificação da Realidade). Tais métricas e critérios foram organizados de forma que possam ser implementados em um sistema para geração automática de relatório sobre a participação dos alunos, que serviria como base para a avaliação do professor. Com isso, a proposta de solução que será apresentada atinge o primeiro processo do ato de avaliar proposto em (LUCKESI, 2011), o Diagnóstico, deixando a intervenção e a avaliação real para os educadores.

Tendo em vista as métricas e os critérios propostos, as mensagens da sessão são extraídas e analisadas para obter as medidas de cada aluno e utilizá-las no cálculo das notas em cada métrica. O objetivo é que, ao final do processo, após calcular as notas individuais para cada métrica, seja gerado um relatório de Performance da Participação, contendo um indicador final da participação dos alunos na sessão de bate-papo (que equivale a um percentual de participação na sessão de cada aluno). Com tal relatório e com o Indicador da Participação, espera-se que os educadores tenham apoio para avaliar os alunos na sessão executada.

A seguir serão descritas as métricas e os critérios adotados nesta pesquisa para a qualificação da participação dos alunos na sessão de bate-papo realizada no estudo de caso.

4.1. Métrica: Tempo de Presença na Sessão

A presença dos participantes é necessária para que a discussão ocorra. Se um aluno não comparecer na sessão de bate-papo, deve receber nota 0% naquela sessão. Se o participante estiver presente durante toda a sessão de bate-papo, deve receber nota 100% de presença. Se o participante entrar em algum ponto ao longo da sessão, deve receber nota de presença proporcional ao tempo que esteve presente na sessão - por exemplo, se o participante entrou nos 15 minutos finais de uma sessão de bate-papo com duração de 60 minutos, então ele deve receber nota 25% de presença.

A estratégia para extração das medidas dos alunos (percentual de tempo de presença de cada aluno) é identificar o horário de entrada e saída do aluno na sessão de bate-papo e comparar com o tempo de início e término da sessão, e assim calcular o percentual em que o aluno esteve presente na sessão.

O critério para qualificação dos participantes nesta métrica é definido como o próprio percentual do tempo em que esteve presente na sessão.

4.2. Métrica: Quantidade de Mensagens

Se um aluno estiver presente ao longo de toda a sessão de bate-papo, mas durante esse período não enviar nenhuma mensagem, sua participação é inadequada – é como se tivesse entrado na sala de bate-papo e fosse fazer qualquer outra coisa. Por outro lado, se o aluno enviar uma enorme quantidade de mensagens, acaba monopolizando a conversa e sua participação também é inadequada. Há, portanto, uma determinada quantidade de mensagens que se espera que o aluno envie ao longo de uma sessão.

A métrica “quantidade de mensagens” corresponde à quantidade de mensagens enviadas pelos participantes durante a sessão de bate-papo, excluindo as mensagens de saudações e despedidas, pois não fazem parte da discussão.

A estratégia para extração das medidas dos alunos (quantidade de mensagens enviadas por cada aluno) é a contagem das mensagens enviadas entre o início e o término da sessão, anunciados pelo tutor. Por exemplo, um aluno que enviou 22 mensagens, sendo uma mensagem de boas vindas (antes do início da sessão) e uma mensagem de despedida (após o término da sessão), totalizou 20 mensagens.

O critério para qualificação dos participantes nesta métrica é definido com base num limite inferior e superior de mensagens enviadas, como estabelecido em (PIMENTEL *et al.*, 2004). O aluno recebe nota máxima caso a quantidade de mensagens que ele enviou esteja dentro do intervalo considerado como desejado, delimitado entre os limites inferior e superior. Se enviar menos mensagens que o limite inferior, considera-se que a participação do aluno está aquém do desejado, e recebe nota menor quanto menos mensagens tiver enviado. Se enviar mais mensagens que o limite superior, considera-se que o aluno está participando inadequadamente monopolizando a discussão, e por isso sua nota é reduzida gradativamente conforme se distancia do limite superior. Por exemplo, considerando uma sessão de 15 alunos e que cada aluno envia, em média, 25 mensagens durante a sessão (como exemplificam os dados levantados no contexto estudado por [ROCHA, 2013]). Desta forma, arbitrariamente pode ser definido que o limite inferior para atribuir nota máxima é de 20 mensagens (80% da média de mensagens da sessão), e que o limite superior para atribuir nota máxima é de 50 mensagens (o dobro da média). Também arbitrariamente, pode-se definir que o extremo inferior para atribuição de nota é de zero mensagem e que o extremo superior para atribuição de nota é de 100 mensagens (o dobro do limite superior). A Figura 5 apresenta a visão gráfica do critério arbitrado para o cálculo da nota na presente métrica.

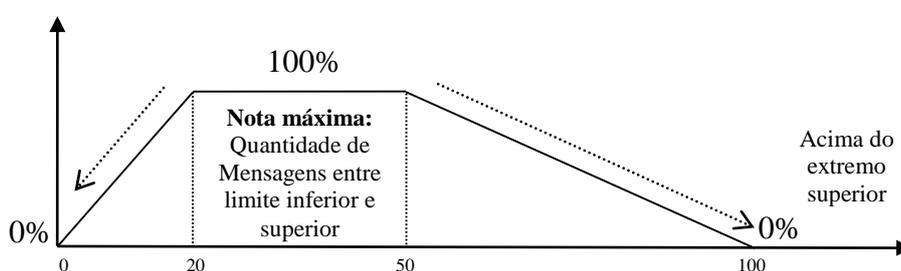


Figura 5. Critério para definição de nota na métrica Quantidade de Mensagens

Neste cenário, caso o aluno envie:

- **De 20 a 50 mensagens:** Nota 100%

Nota máxima, dado que a quantidade de mensagens está dentro dos limites definidos (participação satisfatória).

- **De 0 a 19 mensagens:** Nota diretamente proporcional com a quantidade de mensagens enviadas, sendo que no limite de 20 mensagens recebe nota 100%. Neste caso, se o aluno enviar 12 mensagens recebe nota: $12 * 100\% / 20 = 60\%$.

Ocorre penalização quando são enviadas menos mensagens do que o limite inferior.

- **De 51 a 100 mensagens:** Nota inversamente proporcional à quantidade de mensagens enviadas, sendo que só recebe nota 100% quando enviar no máximo 50 mensagens, e a nota decai até 0% quando enviar 100 mensagens. Por exemplo, se o aluno enviar 65 mensagens, recebe nota: $100\% - [(65-50) * 100\% / 50] = 70\%$.

Ocorre mais penalização quanto mais mensagens o aluno tiver enviado acima do limite superior.

- **Acima de 100 mensagens** (mais que o dobro do limite superior): Nota 0%, visto que excede a quantidade de mensagens aceitável dentro da sessão;

Penalização devido à quantidade de mensagens estar acima do extremo superior (quantidade de mensagens muito alta, sobrecarga de mensagens e monopolização da discussão).

A avaliação da participação em função da quantidade de mensagens, conforme discutido também em outros trabalhos, precisa ser ajustada de acordo com características da sessão. Por exemplo, uma média de 15 mensagens por aluno pode ser considerada alta em uma sessão, e esta mesma quantidade de mensagens pode ser considerada baixa em outro cenário.

4.3. Métrica: Quantidade Média de Caracteres por Mensagem

Se o aluno enviar mensagens muito curtas – por exemplo, apenas com um *emoticon* ou “concordo” –, sem elaborar o texto, sua participação também é considerada inadequada. Por outro lado, mensagens muito longas também são inadequadas para um bate-papo, pois são esperadas mensagens relativamente curtas. Há, portanto, uma determinada quantidade de caracteres que se espera que o aluno envie, em média, em suas mensagens ao longo da sessão.

A métrica “quantidade média de caracteres por mensagem” consiste na média de caracteres enviados por mensagem. O objetivo desta métrica é atribuir nota máxima para os participantes que enviam mensagens de tamanho adequado, e atribuir notas mais baixas para os participantes que enviam mensagens muito curtas (com poucos caracteres), e também atribuir notas mais baixas para os que escrevem mensagens excessivamente longas (com muito mais caracteres do que a média).

A estratégia para extração das medidas dos alunos (média de caracteres por mensagem do aluno) é a contagem da quantidade de caracteres de todas as mensagens enviadas pelo aluno dividido pelo total de mensagens que ele enviou. Por exemplo, um aluno que enviou 10 mensagens que somam um total de 1240 caracteres, teve uma média de 124 caracteres por mensagem.

O critério para qualificação dos participantes nesta métrica também foi baseado na definição de limites inferior e superior para a atribuição da nota máxima, e os extremos inferior e superior para atribuição de nota (como feito na métrica Quantidade de Mensagens). Por exemplo, no estudo apresentado em (ROCHA, 2013), em 142 sessões de bate-papo educacional foram enviadas 52.077 mensagens que totalizaram 3.681.044 caracteres, o que resulta em 70 caracteres por mensagem. A partir desta informação, arbitrariamente pode-se definir que o limite inferior para atribuição da nota máxima é de 35 caracteres (50% da média de caracteres) e o limite superior para atribuição da nota máxima é de 105 caracteres (50% acima da média de caracteres). Também arbitrariamente, pode-se definir que o extremo inferior para atribuição de nota é de zero caractere e que o extremo superior para atribuição de nota é de 210 caracteres (o dobro do limite superior). A Figura 6 apresenta a visão gráfica do critério utilizado para o cálculo da nota na presente métrica.

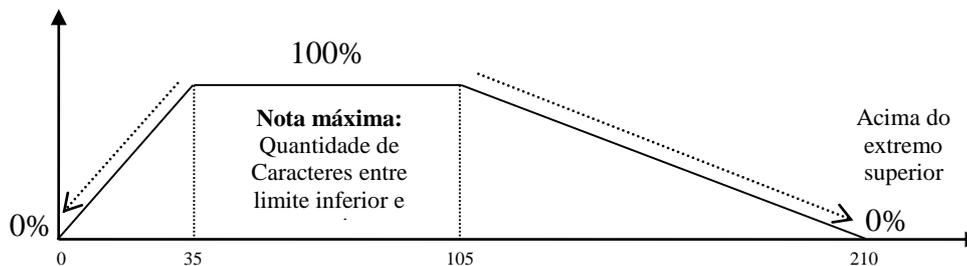


Figura 6. Critério para definição de nota na métrica Quantidade de Caracteres por Mensagem

Neste cenário, caso o aluno envie:

- **De 35 a 105 caracteres por mensagem:** Nota 100%;
Nota máxima, dado que a quantidade de caracteres está dentro dos limites definidos (participação satisfatória, bom tempo para leitura e mensagens bem elaboradas).
- **De 0 a 34 caracteres por mensagem:** Nota diretamente proporcional, de acordo com a quantidade média de caracteres enviados, onde para nenhum caractere a nota será 0% e para 34 caracteres em média por mensagem a nota será de 97%,
Neste caso, ocorre a penalização proporcional para quantidade de caracteres menor que o limite inferior.
- **De 106 a 210 caracteres por mensagem:** Nota inversamente proporcional de acordo com a quantidade média de caracteres enviados, sendo que só recebe nota 100% quando enviar no máximo 105 caracteres por mensagem, e a nota decai até 0% quando enviar 210 caracteres por mensagem.
Penalização devido à quantidade de caracteres estar acima do limite superior.
- **Acima de 210 caracteres em média por mensagens** (acima do extremo superior): Nota 0%, visto que excede a quantidade de caracteres média por mensagem aceitável dentro da sessão, estando acima do extremo superior arbitrado.
Penalização devido à quantidade de caracteres estar acima do extremo superior (quantidade de caracteres muito alta, aumento do tempo de leitura e comprometimento da participação dos demais alunos).

4.4. Métrica: Pertinência das Mensagens

Se a avaliação fosse baseada apenas nos critérios quantitativos (percentual do tempo de presença, quantidade de mensagens enviadas e a média de caracteres por mensagem), o aluno poderia enviar mensagens sem muita importância para a discussão apenas para tirar uma boa nota. Portanto, é preciso de alguma forma avaliar a qualidade das mensagens enviadas pelo aluno.

A métrica “pertinência de mensagens” consiste no percentual de mensagens que contêm alguma palavra chave da discussão. O objetivo desta métrica é caracterizar as mensagens que são pertinentes à discussão.

A estratégia para extração das medidas dos alunos (percentual de mensagens do aluno pertinentes à discussão) é a contagem da quantidade de mensagens que contém alguma das palavras relevantes para o tema (definidas pelo tutor da sessão). Por exemplo, caso o participante tenha enviado 10 mensagens e 4 destas contenham uma das palavras relevantes definidas pelo tutor, seu percentual de pertinência das mensagens será igual a 40%.

O principal ponto de atenção sobre esta métrica refere-se à definição do conjunto de palavras relevantes para o tema da discussão, pois a má definição deste conjunto irá impactar diretamente a nota dos participantes. Como o bate-papo é normalmente dinâmico, definir um conjunto de palavras relevantes apenas no início da sessão e não refiná-lo ao final pode distorcer a avaliação, dado que outros termos importantes podem ser discutidos durante a sessão. Podemos ter, inclusive, variações dos termos (como abreviaturas) utilizadas dentro do contexto, mas não previsíveis sem a posterior avaliação do texto produzido. Portanto, o conjunto de palavras relevantes pode ser gerado em dois passos:

- **Antes da sessão.** O tutor define um conjunto de palavras consideradas relevantes para o tema, elencando os termos que deveriam aparecer na discussão. Por exemplo, se o bate-papo for para discutir “cibercultura e educação”, talvez o tutor possa definir palavras-chave como: cibercultura, cultura, comunicação, educação, aprendizagem, ensino, multireferencialidade, Levy, Lemos, Edmea etc. Devem ser consideradas palavras ligadas ao tema, nomes de autores importantes, técnicas, metodologias e métodos relacionados ao tema da sessão.

- **Após a sessão.** Pode-se extrair os principais termos utilizados nas mensagens enviadas para que sirvam de base para o tutor revisar as palavras-chave que caracterizam aquela sessão. Por exemplo, após a realização da sessão sobre “cibercultura e educação”, dentre as palavras mais citadas nas mensagens dos alunos apareceu o termo “currículo”, que inicialmente o tutor não havia previsto, mas após a sessão pode passar a considerá-la como uma palavra-chave.

O critério para qualificação dos participantes nesta métrica não precisa ser diretamente o percentual de mensagens pertinentes, pois é razoável que o aluno envie algumas mensagens sem conter alguma das palavras-chave. Mas se nenhuma das mensagens do aluno contiver alguma das palavras-chave, então pode ser que o aluno não tenha discutido o assunto esperado. Neste caso, pode-se arbitrar um determinado percentual de mensagens pertinentes a partir do qual o aluno recebe nota máxima neste critério. Por exemplo, suponha que seja arbitrado que a nota máxima seja atribuída ao aluno que enviar 55% das mensagens identificadas como pertinentes. Neste caso, se um aluno tiver enviado apenas 40% das mensagens pertinentes, sua nota seria $40\% \times (100\%/55) = 73\%$.

Com a aplicação deste critério, temos uma avaliação que considera não somente a quantidade de texto produzido, mas também a adequação do texto produzido com o tema em discussão, de acordo com o conteúdo esperado pelo tutor e o resultado da discussão (considerando as palavras frequentes da sessão).

4.5. Nota Final – Indicador da Participação

A estratégia utilizada nesta pesquisa para a qualificação da participação dos alunos na sessão de bate-papo foi a definição de métricas objetivas e critérios que qualificam a participação dos alunos, considerando aspectos variados da participação na sessão. O conjunto de métricas, e os respectivos critérios propostos, devem ser combinados para se caracterizar um valor único, global, aqui intitulado Indicador da Participação.

O Indicador da Participação foi aqui definido como sendo a multiplicação da performance em cada uma das métricas apresentadas neste capítulo. A fórmula utilizada para o cálculo do Indicador da Participação na sessão está descrita na expressão a seguir:

$$\text{Indicador da Participação} = \% \text{Presença} * \% \text{QtdMsg} * \% \text{QtdCaracteres} * \% \text{Pertinência}$$

Por exemplo, se o aluno obteve 100% de aproveitamento em termos de quantidade de mensagens, caracteres e de pertinência das mensagens, mas só esteve presente em 70% da sessão, seu Indicador da Participação será de 70%, dada a ponderação pela presença. Outro exemplo: caso um aluno esteja presente durante 100% da sessão, mas tenha enviado uma quantidade de mensagens insuficiente de tal modo que obtenha 80% da nota relacionada a esse critério; apesar disso, as mensagens que enviou foram suficientemente elaboradas, obtendo nota 100% na quantidade de caracteres; contudo, apenas 40% de suas mensagens continham ao menos uma das palavras-chave da sessão, obtendo nota 73% (conforme exemplo discutido anteriormente); então seu Indicador da Participação é calculado pela fórmula:

$$\text{Indicador da Participação} = 100\% * 80\% * 100\% * 73\% = 58,4\%$$

O objetivo do Indicador da Participação é produzir um valor único para caracterizar e comparar o desempenho dos alunos na sessão. Esse índice não necessariamente precisa ser a nota que o tutor atribui para a participação do aluno naquela sessão, mas pode sim auxiliar o tutor a refletir e avaliar a participação dos alunos. Buscou-se, aqui, ponderar todos os critérios objetivos que julgamos importantes para qualificar a participação dos alunos na sessão de bate-papo. Após elaborarmos esta proposta, foi conduzido estudo de caso exploratório em uma sessão real de bate-papo educacional com alunos de um curso de Graduação em Sistemas de Informação, para avaliar a aceitação das métricas, dos critérios e desse indicador.

5. Avaliação das Métricas e Critérios: um estudo de caso em uma sessão real de bate-papo

As métricas e critérios de qualificação da participação criados nesta pesquisa tiveram como base o estudo de trabalhos relacionados e propostas baseadas na análise de sessões de bate-papo anteriores. Após a definição teórica dos fatores que permitem a qualificação da participação dos alunos, é necessário testar a aplicabilidade e aceitação das métricas e critérios, de forma que seja possível enfrentar problemas e situações reais que podem influenciar nos resultados finais.

O objetivo deste capítulo é descrever como foi planejado e executado o estudo de caso para avaliação das métricas e critérios definidos nesta pesquisa, descrevendo problemas enfrentados e consequentes alterações nos critérios inicialmente elaborados, além de apresentar os resultados obtidos com o estudo de caso aplicado. A seção 5.1 apresentará os detalhes do estudo de caso conduzido e informações sobre os artefatos utilizados, enquanto a seção 5.2 detalha os resultados obtidos com a aplicação do estudo de caso, bem como a avaliação do método de qualificação da participação proposto nesta pesquisa.

5.1 Plano e contexto da avaliação das métricas e critérios

Para a avaliação das métricas objetivas e critérios propostos nesta pesquisa foi conduzido um estudo de caso exploratório, visando testar a geração das notas dos participantes em cada métrica, além da avaliação do funcionamento do modelo proposto dentro do contexto de uma sessão real de bate-papo. O objetivo deste estudo de caso é validar a aplicabilidade e aceitação das métricas e critérios propostos, bem como identificar pontos de melhoria no modelo proposto.

O objetivo da sessão utilizada no estudo de caso foi avaliar a qualidade de sites de internet, sob o ponto de vista da usabilidade. O tutor responsável pela sessão (que se comportou como um moderador) foi o responsável por elaborar pontos de vista sobre os sites e efetuar questionamentos aos participantes para incentivá-los a expor suas

opiniões baseadas nas técnicas conhecidas, com a possibilidade de responderem livremente às questões expostas pelo moderador.

Para a realização da sessão foi usado o serviço de bate-papo do Facebook, dado que algumas atividades do curso já estavam apoiadas na utilização do Facebook para disponibilização de arquivos, entrega de trabalhos e discussões sobre as aulas. Com isso, a iniciativa de utilizar o bate-papo do Facebook como meio para realização da sessão teve aceitação positiva pela turma.

Os alunos foram convidados com aproximadamente uma semana de antecedência a participarem da sessão. Foi informado o horário e o tema que seria discutido na sessão, alertando também que seria executada uma pesquisa após a sessão, relativa à avaliação da participação no bate-papo.

Foram envolvidos na sessão 15 participantes, além do tutor, que atuou como moderador da sessão. Todos os participantes eram alunos de graduação em sistemas de informação, na UNIRIO. As métricas e critérios para qualificação da participação não foram revelados aos participantes antes da sessão, para que não houvesse tentativas de influenciar na avaliação.

A sessão teve duração programada de 1 hora e o tempo real da sessão foi de 58 minutos, desde o início da sessão marcado pelo tutor e o seu término, também indicado pelo tutor. Neste período, os 16 participantes (incluindo o tutor) foram capazes de enviar 345 mensagens, excluindo mensagens de saudação no início e término da sessão. Ao total, foram enviados 24.544 caracteres, o que corresponde a uma média de aproximadamente 71 caracteres por mensagem.

Ao término do bate-papo os participantes foram convidados a responder um questionário de autoavaliação sobre a participação individual na sessão (cada aluno avaliou sua própria participação). O objetivo deste formulário foi identificar a opinião dos alunos sobre sua própria avaliação da participação no bate-papo, que durante a análise dos resultados seria utilizado como base para comparação com as notas atribuídas através das métricas e critérios propostos nesta pesquisa. O objetivo desta comparação é identificar conflitos de opinião e gerar discussão sobre o método de qualificação da participação proposto. O formulário de autoavaliação aplicado continha as seguintes perguntas:

- Em uma escala de 0 a 10, qual nota você atribui à sua participação nesta sessão de bate papo?
- Justifique sua nota
- Informe seu nome

Após as respostas da autoavaliação pelos alunos, encerrou-se a primeira parte do estudo de caso, visto que os alunos completaram todas as tarefas iniciais (participação na sessão e resposta da autoavaliação).

5.1.1 Extração das mensagens da sessão e cálculo das notas em cada métrica

Após a realização da sessão as mensagens enviadas foram extraídas do Facebook e inseridas manualmente em planilha eletrônica, com as colunas “Horário da Mensagem”, “Usuário” e “Mensagem”. Para gerar a planilha, as mensagens foram copiadas da interface de bate-papo do Facebook e coladas em editor de texto, em formato HTML, para então editar o conteúdo colado, visando extrair apenas as informações relevantes para a análise.

A partir da planilha gerada foram avaliadas manualmente as primeiras mensagens da sessão e removidas da planilha aquelas referentes a saudações, cumprimentos e abertura da sessão, mantendo apenas mensagens enviadas após o início formal da sessão. Foram removidas também mensagens de despedida, após o término formal da sessão, marcado pelo tutor.

Com base na planilha gerada foram calculados indicadores individuais para apoio na avaliação das notas dos participantes, como quantidade de mensagens e caracteres enviados por cada participante, horário da primeira e o tempo de permanência na sessão. Estas informações serviram de base para a geração da nota nas métricas Quantidade de Mensagens, Quantidade Média de Caracteres por mensagem e Tempo de Presença na Sessão.

Para estas métricas foram utilizadas fórmulas simples do Microsoft Excel, conforme indicado na Tabela 1.

Tabela 1. Fórmulas para cálculo de indicadores básicos das mensagens

Métrica	Fórmula utilizada
Quantidade de Mensagens	Gerada tabela dinâmica para totalizar a quantidade de mensagens enviadas por participante, após remoção das mensagens irrelevantes.
Quantidade Média de Caracteres por Mensagem	Incluída coluna ao lado das mensagens para contabilizar a quantidade de caracteres, utilizando a fórmula: NÚM.CARACT(célula que contém a mensagem). Gerada tabela dinâmica para totalizar a quantidade de caracteres por participante. Divisão do total de caracteres pelo total de mensagens, para cálculo da média de caracteres por mensagem para cada participante.
Tempo de Presença na Sessão	Anotado manualmente a primeira mensagem de cada participante. Cálculo da diferença entre o horário de término da sessão e a primeira mensagem enviada, para identificar o tempo de participação do aluno, em minutos. Divisão do tempo de participação do aluno pelo tempo total da sessão, em minutos, para obter o percentual do tempo de presença na sessão.

O serviço de bate-papo no Facebook não registra informação importante para o estudo realizado: o horário de entrada e saída de cada participante. Esta informação é importante para calcular o percentual de presença na sessão, conforme o critério estabelecido para esta métrica no capítulo 4.1. Dada esta limitação, a alternativa foi utilizar como entrada na sessão a primeira mensagem enviada pelo participante. Caso a primeira mensagem tenha sido enviada antes do início da sessão declarado pelo tutor, considerou-se que o aluno esteve presente desde o início, caso contrário, considerou-se o horário da primeira mensagem no cálculo do tempo de presença. Como não havia também a informação da saída dos alunos da sessão, foi considerado que todos os alunos que estavam presentes (independente do horário da primeira mensagem) ficaram na sessão até o término.

Para a geração das notas na métrica Pertinência das Mensagens foi implementado um programa para recuperar as mensagens enviadas (e todas as palavras de cada mensagem) para a geração de um gráfico das palavras mais utilizadas, representado por uma *Tag Cloud*, exemplificado na Figura 7. Para esta análise, foi necessária a utilização de métodos básicos de mineração de texto para otimizar a extração das mensagens e gerar o conjunto das palavras mais enviadas nas mensagens dos participantes durante a sessão, desconsiderando termos irrelevantes.



Figura 7. Tag Cloud com palavras mais enviadas

Foram considerados os seguintes passos para aperfeiçoar a lista de palavras mais enviadas:

Passo 1: Recuperação das palavras enviadas ao longo da sessão:

Neste passo foi implementado um programa de computador para a leitura de todas as mensagens enviadas durante a sessão pelos participantes, armazenadas em um banco de dados, e separação das palavras para a geração de um conjunto com as mensagens de cada usuário.

Passo 2: Remoção de *Stop Words* e palavras não desejadas para a análise:

Os *Stop Words* são termos normalmente removidos da análise de processos de mineração de texto, por não terem significado ou relevância dentro do contexto em estudo. Além da remoção de *Stop Words*, na análise realizada neste estudo, para a métrica Pertinência das Mensagens, foram removidas palavras e termos sem valor como verbos, artigos e preposições. O ideal seria a utilização de apenas substantivos relevantes para o tema e remover termos que não façam sentido para o contexto, como alguns enviados em mensagens de saudação ou atenção.

Por exemplo, as frases abaixo contêm palavras (em negrito) que não acrescentarão valor à discussão:

“**Pessoal**, entrei **na sala**”

“**Olá** vamos começar **nossa sessão**”

Passo 3: Geração de Gráfico com as palavras mais utilizadas

Após a remoção das *stop words* e palavras não desejadas, foi gerado um gráfico no formato de *Tag Cloud* (Figura 7) contendo as palavras com mais ocorrência dentro da sessão para análise preliminar dos resultados. A análise inicial apontou necessidade de refinamento da lista de *stop words* para inclusão de outras palavras que foram empregadas com frequência, porém sem ganho para o contexto.

Passo 4: Refinamento das *Stop Words* e palavras não desejadas

Com base no gráfico gerado no passo anterior, foi realizada nova análise das palavras frequentes e a lista de *stop words* e palavras não desejadas foi refinada após algumas execuções, observando a aparição das palavras consideradas como relevantes e removendo termos identificados como desnecessários para a análise. A execução de múltiplas rodadas de análise e refinamento é comum na aplicação de técnicas de mineração de texto e foi a abordagem utilizada nesta pesquisa.

Ao final do processo o gráfico com das palavras frequentes na sessão foi apresentado ao tutor, para que refinasse a lista produzida antes da sessão. Foi identificado que a lista de palavras definidas pelo tutor antes da sessão foi similar à lista de palavras frequentes, geradas através da análise da sessão, após os passos de remoção de *stop words* e palavras não desejadas.

Com a lista de palavras definida, foi utilizado o programa de computador criado para a contagem de mensagens com palavras relevantes para o tema, por participante, com o objetivo de gerar as notas na métrica Pertinência das Mensagens. Durante a análise das notas geradas foi percebido que a utilização apenas do percentual de mensagens relevantes para a nota não seria justo no cenário estudado, pois um aluno atingiria nota máxima apenas se enviasse ao menos uma das palavras selecionadas pelo tutor em todas as mensagens enviadas. Caso fosse considerado apenas o percentual, sem um nivelamento como foi feito para as métricas Quantidade de Mensagens e Caracteres, nenhum aluno teria obtido nota máxima na métrica Pertinência das Mensagens.

O critério para qualificação dos participantes na Pertinência das Mensagens foi então alterado para que:

- caso o percentual de pertinência calculado para o aluno seja igual ou superior a 55%, então a nota será 100%;
- caso o percentual de pertinência calculado para o aluno seja igual ou superior a 35% e menor que 55%, então a nota será 65%; e
- caso o percentual de pertinência calculado para o aluno seja menor que 35%, será considerado o próprio percentual de pertinência das mensagens.

Por fim, após a geração das notas individuais em cada métrica, utilizando os critérios definidos no capítulo anterior, foi possível calcular a nota final de cada aluno,

correspondente ao Indicador da Participação, utilizando a fórmula descrita anteriormente no item 4.5

5.1.2 Geração do Relatório de Performance da Participação

Após o cálculo das notas individuais dos alunos em cada métrica definida e do Indicador da Participação, foi gerado (manualmente) o relatório de Performance da Participação para utilização junto aos alunos na segunda parte do estudo de caso – a avaliação pelos alunos do método utilizado para qualificação da participação, considerando as métricas e critérios propostos nesta pesquisa.

O relatório gerado contém as seguintes áreas:

Tabela 2. Áreas do Relatório de Performance da Participação

Área	Descrição
Introdução	Texto introdutório sobre o Relatório de Performance da Participação, o contexto e quais serão as demais áreas do relatório. O objetivo desta área do relatório foi preparar os alunos as notas em cada métrica e orientá-los sobre como analisar o relatório.
Apresentação das Métricas e explicação sobre os critérios adotados	Para cada métrica foi gerada uma área no relatório que apresentava os seguintes tópicos: 1- Gráfico de Barras com a representação das notas de cada aluno 2- Descrição da métrica e critérios considerados para geração da nota O objetivo destas áreas foi apresentar as notas de cada aluno através de representação gráfica (para facilitar a visualização) e fornecer explicação sobre como a nota foi calculada. O Indicador da Participação, produto das demais notas, considerado como a nota final, foi também apresentado nesta área, com as devidas explicações.
Notas na autoavaliação	Gráfico de barras representando as notas atribuídas por cada aluno na autoavaliação, para que pudessem facilmente comparar com seus o indicadores de participação. O objetivo desta área foi apresentar as notas individuais no mesmo formato do Indicador da Participação, para facilitar a discussão das diferenças entre a autoavaliação e a qualificação da participação utilizando as métricas e critérios propostos nesta pesquisa.

O relatório gerado foi disponibilizado aos alunos através do Facebook para que todos os participantes tivessem acesso às informações. Neste ponto, a tarefa dos alunos foi analisar o relatório e verificar sua coerência, de acordo com a opinião de cada um. Junto com o relatório foi disponibilizado um questionário para avaliação do relatório, pelos alunos.

5.1.3 Questionário de Avaliação do Relatório de Performance da Participação

O objetivo do questionário aplicado foi coletar a opinião dos alunos sobre a concordância com as métricas e critérios estabelecidos para a qualificação da participação na sessão. Indiretamente, objetivou-se também verificar se o formato utilizado para o relatório estava adequado de fácil entendimento pelos alunos. O questionário foi organizado com a seguinte estrutura:

Tabela 3. Estrutura do questionário de avaliação do Relatório de Performance da Participação

Área	Descrição
Identificação	Campo para que o aluno informe seu nome
Avaliação das métricas e critérios	<p>Para cada métrica foram feitas as seguintes perguntas:</p> <ol style="list-style-type: none">1- Você entendeu o critério de avaliação (considere o texto explicativo no relatório)? Respostas possíveis: Não entendi, Entendi parcialmente, Entendi totalmente.2- Você concorda com o critério utilizado para avaliar a nota relativa à quantidade de mensagens? Respostas possíveis: Não concordo, Concordo parcialmente, Concordo totalmente.3- Você concorda com a sua nota de participação pela Quantidade de Mensagens? Respostas possíveis: Não concordo, Concordo parcialmente, Concordo totalmente.4- O que você mudaria no critério de avaliação da Adequação à Quantidade de Mensagens enviadas apresentado no relatório? Campo para resposta em texto livre. <p>O Indicador da Participação também teve área com as perguntas acima, na mesma organização das demais métricas.</p>
Comentários Gerais	<p>A pergunta final, de contexto geral, foi:</p> <p>“O que você mudaria no relatório de Performance da Participação apresentado, seus critérios e/ou formato de avaliação?”</p> <p>Objetivo desta pergunta foi criar um espaço onde os respondentes pudessem sugerir mudanças, enviar críticas a qualquer ponto do relatório ou das métricas utilizadas, bem como os critérios de avaliação das métricas.</p>

Após a disponibilização do relatório e preenchimento do questionário pelos alunos, foi iniciada a última parte do estudo, referente à análise das informações enviadas pelos alunos sobre o método de qualificação da participação proposto na presente pesquisa. As próximas sessões apresentarão mais detalhes sobre os resultados obtidos com a aplicação do estudo de caso e conclusões sobre o método de qualificação da participação proposto.

5.2 Resultados e Conclusões sobre a Avaliação utilizando as Métricas

Após a análise dos resultados foi possível chegar a algumas conclusões sobre a aplicação do Relatório de Avaliação da Participação e o método elaborado nesta pesquisa. Este capítulo apresenta o resultado da análise dos dados gerados na pesquisa, com a avaliação dos alunos quanto ao método utilizado e seus comentários.

Até este ponto da dissertação, fizemos uma separação clara do que chamamos de métrica e critério. A partir deste ponto, neste capítulo, estamos considerando que a avaliação dos alunos corresponde ao conjunto “Métrica + Critério”. Portanto, iremos nos referir apenas à métrica, considerando que o conceito aqui engloba a métrica adotada e o critério utilizado para qualificar a participação em tal métrica.

5.2.1 Quantidade Média de Caracteres por Mensagem – métrica mais aceita considerando a obtida nota

A métrica de avaliação pela quantidade média de caracteres obteve o maior grau de aceitação pelos participantes da sessão. Esta foi também a métrica de mais fácil entendimento, de acordo com a avaliação dos alunos. Após a avaliação do Relatório, 87% dos participantes informaram que entenderam totalmente a métrica de avaliação e 13% entenderam parcialmente, 67% dos participantes concordaram totalmente com a métrica e 87% concordaram totalmente com a nota que obtiveram nesta métrica.

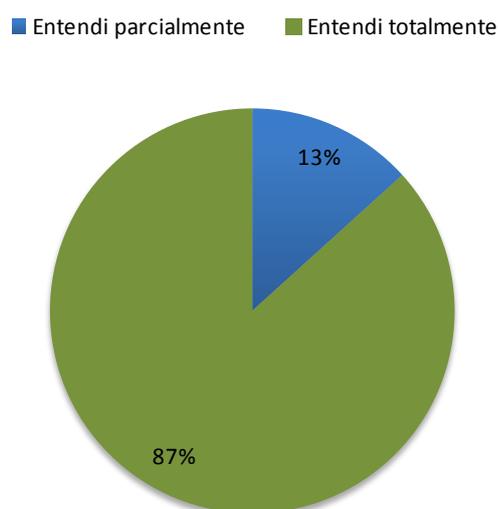


Figura 8. Entendimento da Métrica

■ Concordo parcialmente ■ Concordo totalmente ■ Não concordo

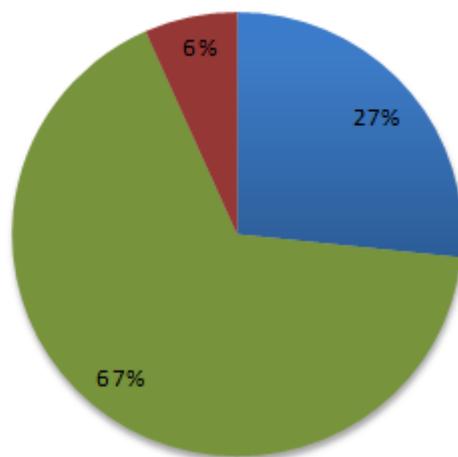


Figura 9. Concordância com a métrica

■ Concordo totalmente ■ Não concorda ■ Concordo parcialmente

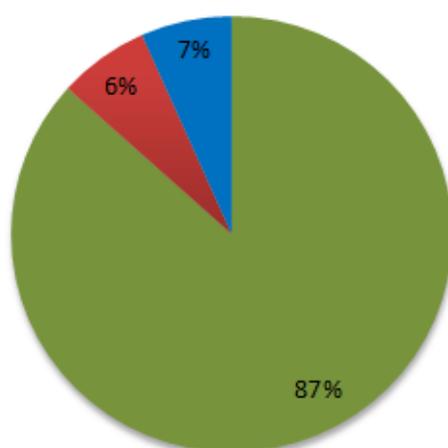


Figura 10. Concordância com a nota obtida na métrica

Foi observado que as notas desta métrica foram máximas (100%) para 12 dos 16 participantes, sendo este um dos pontos que pode ter influenciado na aceitação da nota atribuída. Dos 4 participantes que ficaram abaixo da nota máxima, o que teve a nota mais baixa nesta métrica informou ter entendido o método utilizado, mas não concordar tanto com o método quanto com a sua nota, porém não fez comentários sobre possíveis melhorias. Os 3 demais participantes concordaram totalmente ou parcialmente tanto com a métrica, quanto com a nota obtida. Neste caso, fica claro que o método foi negado pelo participante que não obteve nota satisfatória e bastante aceito pelos demais.

Algumas contribuições interessantes foram relatadas, por exemplo, a participante “Renata”¹ acrescenta comentário sobre possível melhoria ao critério, transcrito integralmente a seguir:

Estenderia um pouco além de 105 o limite para nota 100%, porque, dependendo do que deve ser dito, são precisos mais caracteres do que estes formular uma boa resposta. Mas não muito mais, já que de fato um texto longo interfere no bom andamento desse tipo de debate.

Outros participantes concordaram que o envio de muitos caracteres pode ser prejudicial para o entendimento do assunto e da discussão, conforme o trecho abaixo do participante “José Vinícius”:

Falar demais tira a atenção de qualquer um, basicamente dizendo.

Um ponto de atenção para a métrica foi bem relatado pela participante “Mariella”, sobre a dificuldade em medir a quantidade de caracteres ideal para a sessão. É importante avaliar e refinar a quantidade de caracteres esperada por sessão, visto que mensagens com muitos caracteres podem revelar usuários com textos bem escritos e concisos. A contribuição da participante apresenta essa preocupação:

Quando levamos em conta o número de caracteres por mensagem podemos estar deixando de levar em conta se a mensagem é prolixa ou enxuta. É difícil estimar um número padrão de caracteres para uma mensagem se as pessoas desenvolvem um assunto de forma diferente, ainda mais se tratando de pessoas com pouca experiência no assunto (estudantes iniciantes no assunto). Bem ou mal, as pessoas avaliadas podem precisar de mais palavras para abordar um assunto, não sendo, necessariamente prolixa e sim, mais completa.

De acordo com o resultado desta métrica e os comentários enviados pelos participantes, entendemos que métrica de avaliação pela Quantidade Média de Caracteres é pertinente para a qualificação da participação em sessões de bate-papo e possui boa aceitação pelos envolvidos.

¹ Os nomes dos participantes foram alterados por nomes fictícios, para não revelar a identidade dos alunos.

5.2.2 Tempo de Presença na Sessão – melhor entendimento e maior concordância com a métrica

A métrica Tempo de Presença na Sessão foi a que apresentou maior percentual de entendimento pelos participantes, onde 100% dos envolvidos reportaram ter entendido totalmente. Considerando a concordância com o método, 87% dos participantes informaram que concordaram totalmente com a métrica, sendo o maior percentual para todas as métricas neste quesito.

A concordância com a métrica indica que a avaliação da permanência na sessão deve continuar sendo considerada na qualificação da participação, reduzindo a nota de participantes que não estejam presentes desde o início da sessão. Os gráficos a seguir apresentam o resumo das respostas dos participantes para esta métrica.

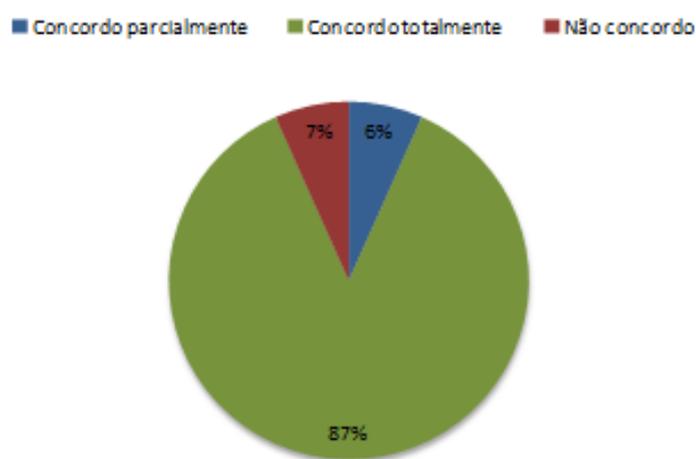


Figura 11 – Concordância com a métrica

A concordância com a métrica foi ressaltada por alguns participantes, como “Josiel Quintanilha” que comentou:

(...) O critério está correto e avalia que as pessoas que chegaram atrasado perderam ponto.

e a participante “Deise Itália”, que informou “Foi bem determinado o parâmetro utilizado.”

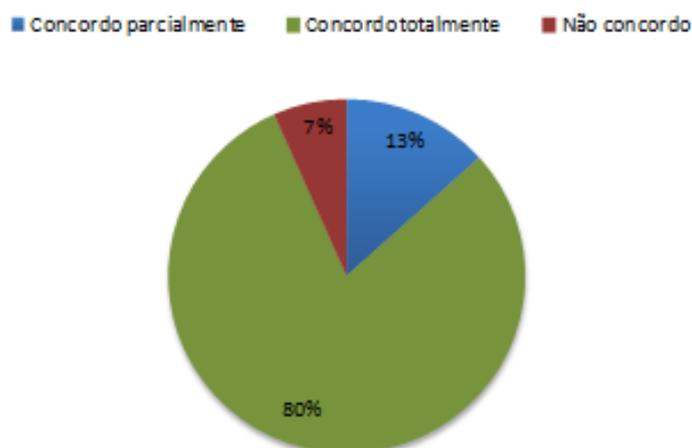


Figura 12 – Concordância com a nota obtida na métrica

Apenas o participante “Ariel Luis” informou não concordar tanto com a métrica quanto com sua nota nesta métrica, mas não informou comentários que justifiquem sua posição. Durante a análise foi identificado que este mesmo participante obteve nota máxima no critério e teve presente durante toda a sessão, logo sua discordância não está relacionada à nota obtida, como em casos anteriores.

Os dois participantes que obtiveram nota baixa na avaliação pela métrica de Presença na Sessão informaram que concordaram parcialmente com a métrica. O participante “José Vinicius” participou de apenas 28% da sessão e “Alex Brasil” participou de 45%. Mesmo com as notas baixas e penalização na nota final, os participantes não discordaram da métrica (concordaram parcialmente), indicando que entendem o sentido da avaliação. Os participantes enviaram contribuições sobre o critério.

José Vinicius:

Não estava disponível na hora que começou formalmente, tinha voltado só após resolver assuntos pessoais em casa ao computador. Não acho que justo que o tempo em que comecei a responder seja mais importante do que quanto me esforcei para ser pertinente ao trabalho.

Alex Brasil:

Minha permanência, além de outros critérios, foi afetada por problemas técnicos, tanto por minha parte como pela parte do organizador. Acredito que esse quesito deveria ter peso menor do que os outros.

Através das mensagens acima, enviadas pelos participantes com baixa nota na avaliação desta métrica, percebe-se que os envolvidos assumem sua ausência em parte da sessão, mas justificam sua concordância parcial. Uma das sugestões que poderia ser analisada e definida em estudos futuros é a redução do peso desta métrica, mas considerando o cenário geral não parece necessário o investimento de tempo para ajustá-la.

Concluimos que o tempo de presença na sessão deve continuar sendo usada para a qualificação da participação na sessão, visto que, assim como em salas de aula tradicionais, a presença dos alunos é importante para a discussão dos temas.

5.2.3 Pertinência das Mensagens – métrica com menor aceitação

As métricas e critérios aplicados consideraram a avaliação quantitativa da sessão, como ocorre com a métrica referente à quantidade de mensagens enviadas. A métrica inserida para a avaliação qualitativa da participação foi a Pertinência das Mensagens, através da identificação do percentual de mensagens enviadas que possuem ao menos uma das palavras-chave definidas como relevantes para a discussão.

Esta é a métrica onde os participantes alcançaram menor nota dentre as outras aplicadas (exceto pelo Indicador da Participação, que será avaliado em separado, visto que é influenciado pelas demais métricas). Aproximadamente 47% dos participantes (7 participantes) conseguiram alcançar a nota máxima nesta métrica. Este volume foi considerado baixo, comparando com as demais métricas, que tiveram em média 73% de participantes com a nota máxima. Outros 40% dos participantes (6 participantes) ficaram com nota igual a 65% e 13% dos participantes (2 participantes) obtiveram nota igual a 30% de aproveitamento.

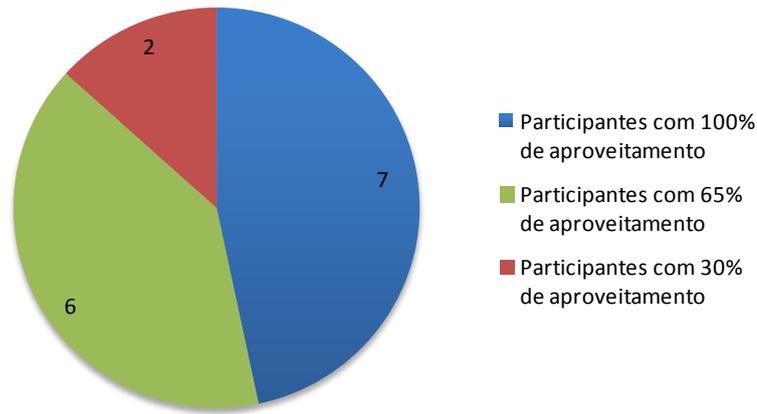


Figura 13 - Distribuição das notas na Métrica Pertinência das Mensagens

As notas baixas nesta métrica tendem a gerar discordâncias sobre o critério adotado para avaliar, visto que os participantes podem considerar a avaliação injusta. Por consequência, este foi o critério com menor aceitação pelos participantes, apesar de ter obtido bom entendimento (87% entenderam totalmente a métrica).

Apenas 40% dos participantes concordaram totalmente com a métrica e 47% concordaram com a nota obtida. Esta foi a métrica mais polêmica adotada para a avaliação e também a que deve gerar maior esforço para ajustes e melhorias para aplicação em outros estudos de caso. Os gráficos abaixo apresentam um resumo das informações enviadas pelos participantes.

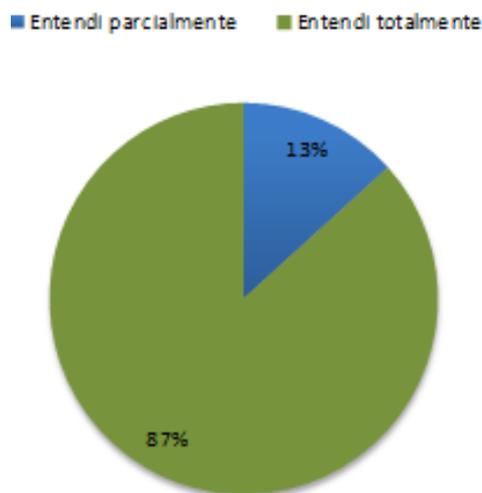


Figura 14. Entendimento da Métrica Pertinência das Mensagens

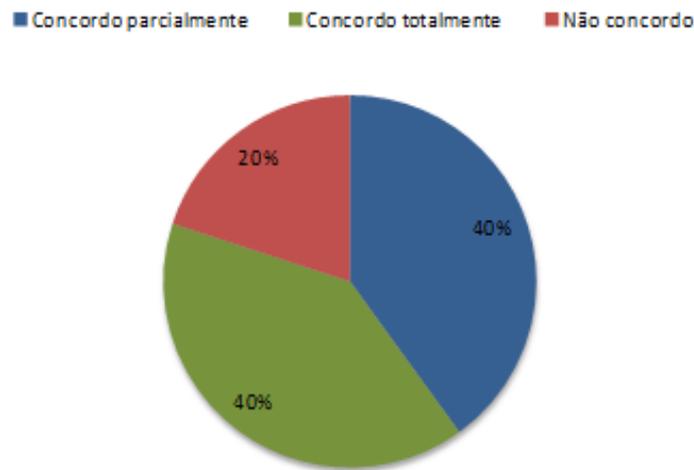


Figura 15 - Concordância com a métrica Pertinência das Mensagens

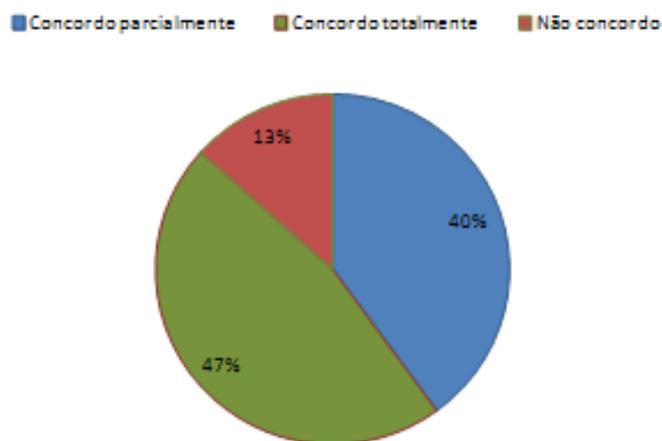


Figura 16. Concordância com a nota obtida na métrica Pertinência das Mensagens

Um dos pontos demonstrados pelos gráficos é que o entendimento do método teve um percentual consideravelmente positivo, dado que é a métrica que mais envolveu análise das informações. Com o bom entendimento da métrica, a discordância reportada pelos participantes realmente refere-se à opinião dos participantes sobre mesma. Dos participantes, 3 informaram que não concordam com métrica e 2 informaram que não concordaram com a nota obtida.

O participante “Ariel Luis” obteve boa nota, mas ainda assim relatou que não concorda com a métrica e com a sua nota. No entanto, este participante não enviou comentários sobre sua discordância. Dado que o participante não concorda, mesmo com a boa nota obtida, seria interessante recuperar mais informações sobre a discordância.

Os demais participantes que discordam da métrica foram “Deise Itália” que obteve o percentual de 30% (o mais baixo da turma, junto com o participante “Yago

Borbato”, que informou concordar totalmente com a métrica e parcialmente com a nota obtida) e “Alex Brasil” que obteve nota 65% (o mesmo percentual de diversos outros participantes, que concordaram parcialmente tanto com a métrica quanto com a nota obtida).

Vale ressaltar que dos dois participantes com nota igual a 30% nesta métrica (nota mais baixa), apenas um relatou não concordar com a métrica. O outro participante informou que concorda com a métrica e concorda parcialmente com a nota obtida. Dos seis participantes que obtiveram nota 65%, apenas um informou não concordar com a métrica e nota obtida e os seis outros concordaram parcialmente ou totalmente com o tanto com a métrica quanto com a nota obtida.

A participante “Deise Itália”, que obteve nota 30% na métrica enviou o comentário abaixo:

O critério é falho, pois usei diferentes tipos de palavras que não foram consideradas na métrica e fiquei com percentual muito menor que de muitos que não falaram nada novo, somente repetiram o que já havia sido dito. Se eu tivesse mandando todas as mensagens com a palavra site, texto, design, nelas, mesmo não tendo compatibilidade com o tema, minha nota seria maior. Mas usei palavras diferentes e fiquei com nota menor do que quem só mandou algumas poucas mensagens com ideias copiadas ou já ditas ou repetindo a frase anterior.

A reclamação enviada pela participante ressalta a necessidade de refinar o processo de seleção de palavras-chave a serem consideradas na análise. O posicionamento da participante é forte, pois não concordou com a métrica e também com a nota obtida. É fato que a métrica não consegue avaliar e penalizar mensagens repetidas que contenham palavras chave, visto que a análise é baseada na ocorrência de ao menos uma das palavras-chave na mensagem enviada.

Para melhorar a assertividade do critério, poderia ser considerado um número maior de palavras-chave, mas deve ser analisado, visto que o aumento das palavras pode levar a avaliações muito positivas, mesmo em casos onde as mensagens não sejam realmente relevantes para o tema.

O participante “Alex Brasil” enviou contribuição simples sobre esta métrica:

Adicionaria outros critérios além de palavras-chave.

A sugestão do participante foi vaga para a análise, visto que não foram relatadas mais informações sobre os critérios que deveriam ser adicionados. A posição deste usuário é menos forte que a anterior, visto que concordou ao menos parcialmente com a nota obtida.

O critério de Pertinência das Mensagens foi ajustado durante a análise dos resultados, conforme apresentado em capítulo anterior, mas as notas geradas para esta métrica apresentam necessidade de maior refinamento e entendimento do formato que pode ser aplicado para esta avaliação. A discordância dos usuários é um importante ponto de entrada para a execução de melhorias na métrica, que foi a mais questionada pelos participantes.

5.2.4 Análise da nota final – o Indicador da Participação

A avaliação das métricas foi realizada em separado, observando os resultados obtidos em cada métrica para qualificação da participação. O produto das notas percentuais de obtidas em cada métrica individual deu origem ao Indicador da Participação. A multiplicação dos percentuais faz ajustes sucessivos na nota considerando o percentual obtido. Com isso, a nota final é balanceada de acordo com as demais.

A média da turma ficou abaixo do esperado – 55%, considerando que de forma geral a sessão foi considerada produtiva e com bom nível de participação pelo tutor (em pergunta realizada após a sessão). Dos 15 participantes (removendo o tutor da sessão), apenas 1 alcançou a nota máxima e 6 (40% ao total) obtiveram nota inferior a 45% de no Indicador da Participação. A quantidade alta (quase metade da turma) de participantes com notas finais baixas foi o principal causador da baixa média da turma na sessão. O gráfico abaixo apresenta a distribuição das notas finais.

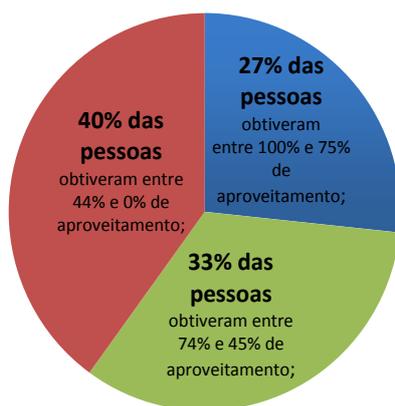


Figura 17. Distribuição das notas no indicador da Participação

Dentre os participantes no grupo de usuários com notas inferiores a 45% no Indicador da Participação, praticamente todos (5 dos 6 participantes) informaram na autoavaliação que acreditavam ter alcançado nota final entre 90% e 100%. Se considerarmos que os participantes tendem a se auto avaliarem com notas superiores a que realmente deveriam receber, ainda assim as notas no Indicador da Participação estariam abaixo do desejado.

Os 6 participantes com Indicador da Participação inferior a 45% estão melhor detalhados abaixo, considerando cada critério individual.

- Alex Brasil (pior caso):

Tabela 4. Análise de Notas do participante Alex Brasil

Critério	% de aproveitamento (nota)	Observações
Quantidade de Mensagens	20%	O participante enviou apenas 4 mensagens durante a sessão, o que indicou quantidade muito baixa de participação.
Quantidade Média de Caracteres	74%	A média de caracteres por mensagem ficou abaixo do esperado para a boa participação, mas ficou próximo do mínimo desejado.
Pertinência das Mensagens	65%	Das 4 mensagens enviadas, 2 continham palavras consideradas relevantes, motivo pelo qual alcançou os 65% de aproveitamento.
Tempo de Presença	45%	O participante esteve presente em menos da metade da sessão, o que deve ter influenciado também na quantidade de mensagens enviadas.
Performance da Participação	4%	O produto das notas anteriores representa o Indicador da Participação. Neste cenário, percebemos que a quantidade baixa de presença e mensagens enviadas tornou a nota final do participante bem abaixo do esperado para a turma.

Analisando o questionário respondido pelo participante é possível perceber que teve concordância e bom entendimento das métricas, porém não concordou com a métrica pertinência das mensagens e com a nota final. Informou também que concordou parcialmente com a métrica de presença na sessão. Foram informados os seguintes comentários:

Sobre a métrica Pertinência das Mensagens:

Adicionaria outros critérios além de palavras-chave.

Sobre a métrica de Presença na Sessão:

Minha permanência, além de outros critérios, foi afetada por problemas técnicos, tanto por minha parte como pela parte do organizador. Acredito que esse quesito deveria ter peso menor do que os outros.

Neste ponto, o participante justifica sua ausência na sessão devido a problemas técnicos. No entanto, a entrada tardia na sessão não é negada em seu comentário. A atribuição de peso menor na avaliação poderia beneficiar o participante, neste caso.

Sobre a nota final (Indicador da Participação):

Alteraria o peso de alguns quesitos, como o de permanência.

Percebe-se pelos comentários enviados que o principal ponto de reclamação do participante é a métrica Presença na Sessão. No entanto, mesmo que esta métrica tivesse alcançado os 100% o participante teria recebido uma nota final abaixo da média da turma (caso tivesse participado de toda a sessão, mantendo as outras notas, o participante teria alcançado nota final igual a 10%). A pior nota do participante é relativa à quantidade de mensagens enviadas, que naturalmente foi impactada pelo tempo de presença na sessão.

De qualquer forma, mesmo considerando possíveis problemas, o próprio participante em sua autoavaliação avaliou-se com uma nota baixa (apenas 40%, a menor da turma) e aparentemente este é realmente um caso de baixa participação, mesmo que em menor escala do que o apresentado pelo relatório de participação.

- Yago Borbato:

Tabela 5. Análise de Notas do participante Yago Borbato

Critério	% de aproveitamento (nota)	Observações
Quantidade de Mensagens	50%	O participante enviou apenas 10 mensagens durante a sessão
Quantidade Média de Caracteres	92%	A média de caracteres se aproximou do ideal, tendo ficado em 32 caracteres por mensagem.
Pertinência das Mensagens	30%	3 das 10 mensagens enviadas continham alguma palavra considerada relevante ao tema
Tempo de Presença	100%	Presente em toda a sessão, desde o início.
Performance da Participação	14%	Fica claro que a nota final foi penalizada pela quantidade de mensagens enviadas e pelo critério de pertinência das mensagens.

O participante não concordou com sua nota final, porém informou que concorda com a métrica e critério adotados. Este foi um padrão de suas respostas também para as outras métricas, onde informou concordar totalmente com a métrica e parcialmente com a nota recebida (exceto para o Tempo de Presença, que concordou totalmente com a nota recebida, que foi de 100%).

O participante não enviou comentários sobre pontos de discordância, o que reforça a possibilidade de ter apenas discordado das métricas por não ter obtido boa avaliação.

- José Vinícios:

Tabela 6. Análise de Notas do participante José Vinícios

Critério	% de aproveitamento (nota)	Observações
Quantidade de Mensagens	100%	O participante enviou 22 mensagens durante a sessão
Quantidade Média de Caracteres	100%	A média de mensagens foi dentro do considerado ideal, 96 caracteres por mensagem.
Pertinência das Mensagens	65%	9 das 22 mensagens enviadas continham alguma palavra considerada relevante ao tema
Tempo de Presença	28%	Baixa taxa de presença na sessão prejudicou a nota final
Performance da Participação	18%	Fica claro que a presença na sessão influenciou para a nota final baixa

Através da análise manual das mensagens enviadas pelo participante durante a sessão percebeu-se que a maioria das mensagens estava realmente relacionadas com o assunto em discussão. A participação na sessão também foi satisfatória considerando as métricas: quantidade de mensagens e caracteres. No caso deste participante, o ponto de

maior influência foi a presença na sessão, que foi de apenas 28% do tempo total. Caso tivesse participado de 100% da sessão, a nota final do participante seria de 65% de aproveitamento, acima da média da turma.

Sobre a métrica Pertinência das Mensagens, o participante entendeu totalmente, mas concordou parcialmente com a métrica e nota recebida, informando o seguinte comentário:

É uma questão muito arbitrária. Considero que o que eu disse foi muito pertinente, mas aí fica a critério do avaliador, não meu.

Foi identificado pela análise manual das mensagens enviadas que a maioria delas estavam associadas ao tema em discussão, sendo consideradas pertinentes. A nota obtida nesta métrica foi dentro do padrão esperado, com 65% de aproveitamento.

Sobre a presença na sessão, o participante comentou:

Não estava disponível na hora que começou formalmente, tinha voltado só após resolver assuntos pessoais em casa ao computador. Não acho que justo que o tempo em que comecei a responder seja mais importante do que quanto me esforcei para ser pertinente ao trabalho.

O comentário enviado é relevante. Realmente as contribuições foram válidas para o tema da sessão, no entanto o grupo poderia ter se beneficiado ainda mais caso o participante estivesse presente desde o início da sessão, se mantivesse o mesmo padrão das mensagens. Este é o motivo pelo qual é aplicada a penalização pelo percentual de presença na sessão.

- Deise Itália:

Tabela 7. Análise de Notas da participante Deise Itália

Critério	% de aproveitamento (nota)	Observações
Quantidade de Mensagens	100%	O participante enviou 37 mensagens durante a sessão
Quantidade Média de Caracteres	100%	A média de mensagens foi dentro do considerado ideal, 75 caracteres por mensagem.
Pertinência das Mensagens	30%	11 das 37 mensagens enviadas continham alguma palavra considerada relevante ao tema
Tempo de Presença	100%	Baixa taxa de presença na sessão prejudicou a nota final
Performance da Participação	30%	Fica claro que a pertinência das mensagens influenciou negativamente na nota final

Percebemos neste caso o peso causado na nota final pela métrica Pertinência das Mensagens. A participante alcançou nota máxima em todos os critérios e teve apenas 30% das mensagens enviadas com palavras definidas como relevantes para o tema. A nota final deste participante reforça a necessidade de refinar o conjunto de palavras relevantes utilizadas ou até mesmo o formato de avaliação desta métrica.

A participante, que se auto avaliou com nota final igual a 100%, fez o comentário abaixo sobre a métrica Pertinência das Mensagens, informando também que não concordou com a métrica e critério adotado:

O critério é falho, pois usei diferentes tipos de palavras que não foram consideradas na métrica e fiquei com percentual muito menor que de muitos que não falaram nada novo, somente repetiram o que já havia sido dito. Se eu tivesse mandando todas as mensagens com a palavra site, texto, design, nelas, mesmo não tendo compatibilidade com o tema, minha nota seria maior. Mas usei palavras diferentes e fiquei com nota menor do que quem só mandou algumas poucas mensagens com ideias copiadas ou já ditas ou repetindo a frase anterior.

O comentário enviado pela participante coloca em questão a fragilidade na utilização de palavras-chave para avaliar a pertinência das mensagens. Fica clara a necessidade de definir e aplicar com cuidado o conjunto de palavras-chave para identificar a pertinência das mensagens. Neste caso, mesmo a inclusão de novas palavras provavelmente não melhoraria consideravelmente a avaliação da participante, devido ao vocabulário empregado.

A principal contribuição da participante no questionário foi realizada na área de comentários livres (sobre o que o participante acredita que deveria ser alterado no método de qualificação da participação) sobre o Relatório de Performance da Participação, transcrito na íntegra abaixo:

Mudaria como são calculados os parâmetros anteriores. Como eu disse anteriormente ele não faz jus às pessoas que tem amplo vocabulário, ou que respondeu aos questionamentos em uma só postagem, reduzindo o número de mensagens enviadas, mas não o

conteúdo. O método é falho. Não condiz com a realidade presenciada na sala durante o debate. Pessoas que copiam e colam mensagens anteriores, para dar continuidade ao pensamento ou somente concordar acabam ficando com notas maiores que das pessoas que realmente se prepararam e viram defeitos no site, assim como coisas bem feitas. Faltaram diversas palavras utilizadas como busca, barra de busca, dificuldade, html, código, fotos, layout, estilos, estilo, regra, defeito, formatação entre muitos outros utilizados para descrever os problemas.

Por esses motivos, acredito que há muito o que melhorar nos parâmetros utilizados para avaliar cada critério.

No caso da participante Deise Itália os comentários são coerentes com a realidade e o único critério que ocasionou a nota final baixa foi a taxa de pertinência das mensagens, igual a 30%. Como este foi a métrica com maior número de discordância entre os participantes (20% dos participantes não concordaram e 13% não concordaram com sua nota nesta métrica), podemos concluir que realmente será necessário aplicar alguns ajustes e, principalmente, melhorar o processo de escolha das palavras-chave que serão consideradas.

- Lucio Camilo e Ariel Luis

O cenário dos participantes Lucio Camilo e Ariel Luis é semelhante na nota final do Indicador da Participação (com 41% e 44%, respectivamente) e nos critérios onde foram penalizados. Ambos tiveram notas baixas nas métricas Quantidade de Mensagens e Caracteres. A análise das mensagens enviadas revela que realmente a quantidade de mensagens enviadas e a qualidade na formação das mesmas foi inferior aos demais participantes com notas melhores.

Um ponto interessante sobre estes participantes é que a nota final não foi ainda menor devido à boa avaliação nas métricas Tempo de Presença na sessão e Pertinência das Mensagens, onde ambos alcançaram a pontuação máxima. Sobre a métrica Pertinência das Mensagens, praticamente todas as mensagens enviadas continham palavras-chave para o tema e este ponto foi importante para a nota final não fosse ainda menor.

O participante Ariel Luis não concordou com a maioria das métricas, porém concordou com sua nota final (Indicador da Participação) atribuída no relatório de avaliação e fez uma contribuição geral sobre o método:

Não acredito que este seja um bom método de avaliação, pois muitas notas de alunos no relatório não condizem com o que observei no debate. Acho que o método tem seu mérito para classificar uma participação como ativa, mas não como construtiva.

É fato que a avaliação da participação como construtiva depende de outros fatores, como refinamento da análise qualitativa das mensagens e de avaliação humana manual para avaliar se os participantes realmente conhecem mais sobre o tema após a sessão. O objetivo do método está ligado à qualificação da participação como ativa na sessão e não ao conhecimento adquirido com a sessão.

Por fim, analisando de forma geral os comentários e detalhes dos participantes com nota final inferior a 40% no Indicador da Participação é possível perceber que alguns participantes realmente tiveram nota compatível com sua contribuição para a sessão, mas em alguns casos a avaliação não parece justa. O principal caso onde a nota final não parece coerente com a contribuição dada na sessão refere-se à participante Deise Itália, devido ao critério de Pertinência das Mensagens.

Para melhor aproveitamento do método de qualificação da participação, a métrica Pertinência das Mensagens precisaria ser refinada para incluir maior quantidade de palavras-chave e/ou utilizar um peso diferenciado na avaliação. Esta foi a métrica que mais influenciou notas finais baixas e também foi a métrica com maior número de discordâncias, entre os respondentes do questionário.

5.2.5 Entendimento das métricas individuais e da nota final (Indicador da Participação)

Analisando as respostas enviadas pelos participantes foi possível perceber que a explicação fornecida para cada métrica no Relatório de Performance da Participação foi suficiente para o entendimento do formato de avaliação e regras utilizadas. O maior problema foi apresentado no entendimento da Nota Final (Indicador da Participação),

onde apenas 53% dos participantes informaram ter entendido totalmente a métrica e o critério e 13% informaram não ter entendido.

A Nota Final (Indicador da Participação) foi a única métrica onde algum participante informou não ter entendido as regras. Não foram enviados comentários específicos relacionados ao não entendimento da Nota Final, porém foi aparentemente causado pela multiplicação dos percentuais obtidos nas demais métricas. Em notas compostas de diversos critérios, é mais comum utilizar a média aritmética ou ponderada entre as notas intermediárias. Após o estudo, percebeu-se inclusive que a utilização de média aritmética das notas teria melhorado o pior caso de nota final, de 4% para 51% (ainda baixo, porém aceitável).

A métrica Tempo de Presença na sessão teve 100% dos participantes informando que entenderam totalmente a métrica e regras adotadas e foi o item com maior quantidade de respostas informando entendimento total, seguida pelas métricas Pertinência das Mensagens (87%), Quantidade Média de Caracteres (87%) e Quantidade de Mensagens (67%).

A Figura 18 apresenta um resumo do entendimento dos participantes sobre as métricas, somando as respostas para todas as métricas. Como pode ser visto, um número reduzido de participantes informou não ter entendido as métricas e 79% entenderam totalmente. Somando o entendimento total e parcial, temos 98% dos participantes informando que tiveram entendimento, mesmo que parcial, das métricas, fator que consideramos relevante para a pesquisa.

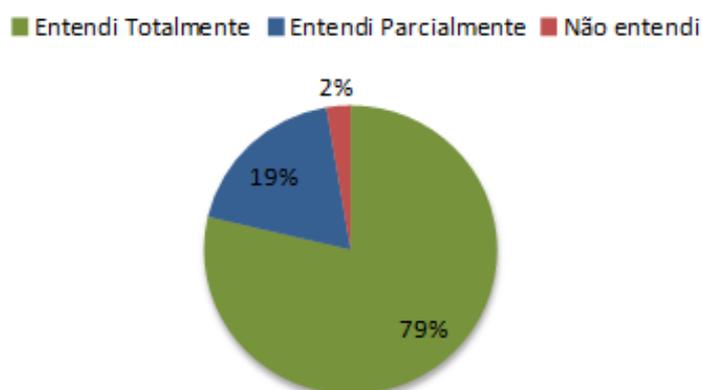


Figura 18. Resumo do Entendimento das Métricas

A fórmula utilizada para geração do Indicador da Participação resultou no pior índice de entendimento pelos participantes, comparando com as demais métricas, e o único que obteve respostas “Não Entendi” (por dois participantes). O Indicador da Participação teve também 5 participantes (34%) informando que entenderam parcialmente e 8 participantes (53%) informando que entenderam totalmente o critério e cálculo realizado.

Dado que o critério e o método utilizado para geração da nota final consistem na multiplicação dos percentuais obtidos nas demais métricas, era esperado que o percentual de entendimento fosse alto. A baixa taxa de entendimento releva que a explicação e o formato de apresentação da nota final precisam ser melhorados para facilitar o entendimento dos participantes.

5.2.6 Concordância com as Métricas e Método para cálculo do Indicador da Participação

Os resultados do questionário respondido pelos participantes para verificação de opiniões sobre o relatório de performance da participação mostram que de modo geral os participantes concordaram com as métricas individuais definidas para a qualificação da participação. O percentual mais baixo de discordância ocorreu com a métrica Quantidade Média de Caracteres (6% responderam “Não Concordo”, equivalente a 1 participante) e a métrica Tempo de Presença na sessão obteve a maior concordância (87% responderam “Concordo Totalmente”, 13 participantes ao total).

Analisando o panorama geral, unindo todas as métricas e a nota final, é possível perceber que mais da metade dos participantes concordaram com as métricas. Se considerarmos a concordância parcial e total, temos que 85% dos participantes tiveram no mínimo concordância parcial e apenas 15% dos participantes não concordaram com as métricas. O gráfico abaixo apresenta a distribuição das opiniões sobre concordância com as métricas e o destaque para a discordância.

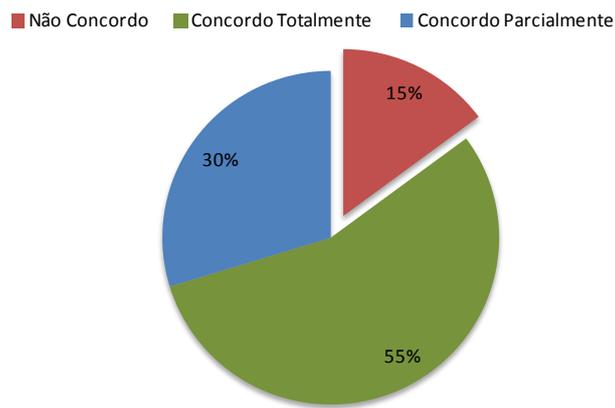


Figura 19 - Concordância com as métricas

Dentro dos 15% de discordância, conforme apresentado na Figura 19, observou-se que a maioria ocorreu na nota final (Indicador da Participação, com 5 discordâncias no total) e na métrica Pertinência das Mensagens (com, 3 discordâncias). Somadas, as duas métricas representam 73% das discordâncias, mais que o dobro do somatório da discordância nas demais métricas. A Figura 20, a seguir, apresenta a distribuição das discordâncias entre as métricas, incluindo a nota final.

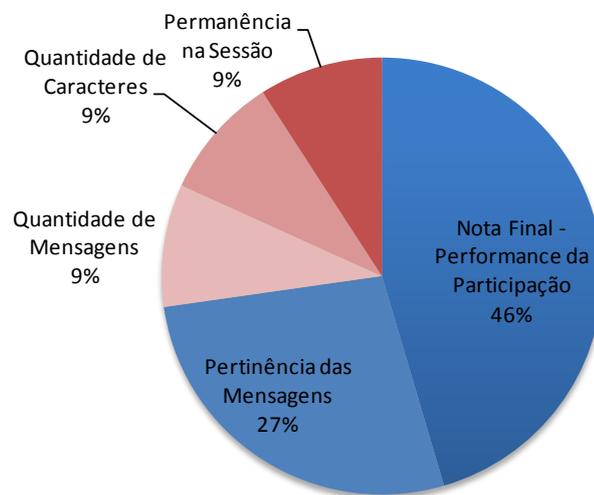


Figura 20 - Distribuição da Discordância entre as métricas

Através da análise das discordâncias apresentadas, fica clara a necessidade de melhorar o método de cálculo do Indicador da Participação e o formato de avaliação da Pertinência das Mensagens, que receberam críticas e comentários expostos em sessões anteriores.

5.2.7 Concordância com as notas calculadas pela solução proposta

Além de avaliar as métricas individualmente, os participantes foram questionados sobre a concordância com a nota gerada pelo método de qualificação da participação proposto nesta pesquisa. O objetivo foi verificar se os participantes concordaram com suas notas, mediante a apresentação dos critérios utilizados, mesmo em casos onde a autoavaliação tivesse sido mais positiva.

Somando o resultado das métricas individuais, os participantes em geral concordaram, mesmo que parcialmente, com as notas obtidas através da solução proposta. Analisando a concordância com as notas, incluindo os participantes que concordaram parcialmente e totalmente com as notas obtidas, chegamos ao percentual de aceitação de 83% da nota atribuída pela solução proposta, conforme apresentado na Figura 21, destacando o percentual de participantes que não concordaram.

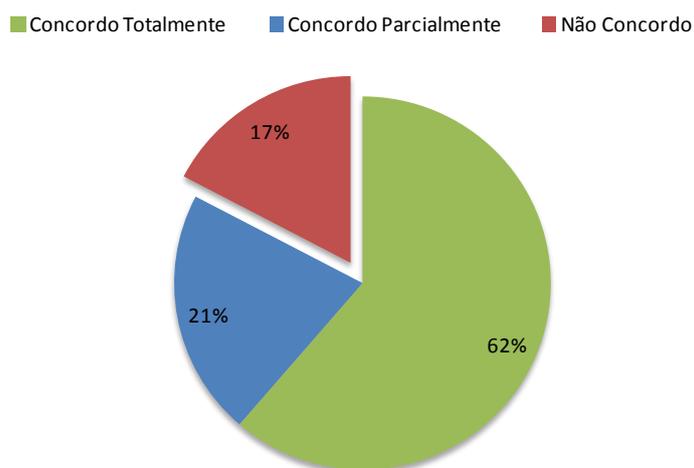


Figura 21 – Concordância com as notas obtidas

As métricas individuais onde foram verificadas maiores discordâncias com relação às notas obtidas foram Quantidade Média de Mensagens e Pertinência das Mensagens, ambas com 2 participantes (13% do total de 15) informando não concordarem com a nota. A nota final (Indicador da Participação) obteve 7 respostas “Não Concordo” (47% dos 15 participantes).

Com relação à Pertinência das Mensagens, três participantes informaram que não concordaram com o critério, um deles informou concordar parcialmente com a nota obtida. Tal participante informou em campo para texto livre que “adicionaria outros

critérios, além das palavras-chave”, confirmando que não discorda totalmente do formato.

Com relação à nota final, todos os participantes que não concordaram com a regra para cálculo também informaram discordância com a nota obtida (5 participantes no total). Dos 7 participantes que informaram não concordar com a nota final obtida, apenas 5 não concordaram com o critério para o cálculo, visto que 1 participante informou concordar parcialmente e outro concordar totalmente.

A Tabela 8 mostra a correlação entre as respostas dos participantes para as duas métricas com menor aceitação pelos participantes, onde 4 dos 5 participantes que informaram concordar totalmente com a nota final obtida também informaram concordar com o critério e com a nota obtida na métrica Pertinência das Mensagens.

Tabela 8. Correlação entre respostas para métricas com menor aceitação

Você concorda com o critério utilizado para avaliar a métrica Pertinência das Mensagens (utilizando palavras-chave)?	Você concorda com a sua nota de participação pela Pertinência das Mensagens?	Você concorda com o critério utilizado para a avaliação final da Performance da Participação na sessão?	Você concorda com a sua nota final, referente ao Indicador da Participação?
Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente
Concordo parcialmente	Concordo totalmente	Concordo totalmente	Concordo totalmente
Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concordo totalmente
Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente
Não concordo	Não concordo	(Não Respondeu)	Concordo totalmente

De acordo com o quadro, todos os participantes que concordaram totalmente com sua nota final concordaram total ou parcialmente (no caso de um dos participantes) com todas as demais notas individuais.

Dos 3 participantes que concordaram parcialmente com sua nota final, 1 concordou totalmente com todas as notas individuais, 1 concordou totalmente com 3 notas e parcialmente com uma delas e o outro participante concordou totalmente com 3 notas e não concordou com uma delas.

Com isso, podemos concluir que os participantes que concordaram totalmente com a nota final atribuída pelo método concordaram total (a maioria) ou parcialmente com todas as demais notas. Os participantes que concordaram parcialmente com a nota final concordaram parcialmente com no máximo duas outras notas ou não concordaram com no máximo uma nota.

Por fim, dos participantes que não concordaram com sua nota final (7 participantes ao total) é possível perceber que:

- Não concordaram com o critério adotado para o cálculo da nota final (5 dos 7 participantes). Logo, se vimos que se o usuário não concordou com o critério, também não concordou com a nota final em todos os casos; ou
- Não concordaram com ao menos uma das notas individuais (um dos casos); ou
- Concordaram parcialmente com ao menos duas notas individuais;

As análises feitas acima sobre os dados apresentados na Tabela 8 excluem as respostas enviadas pelo participante Ariel Luis, visto que a concordância com a nota final parece incoerente, visto que discordou de todos os demais critérios e notas atribuídas. Este comportamento é diferente da maioria dos usuários.

Dos 7 participantes que não concordaram com a nota final atribuída pela solução proposta, a maioria (4 dos 7 participantes) está no grupo de participantes que receberam nota final inferior a 45%. Caso consideremos a nota inferior a 50%, 5 dos 7 participantes que não concordaram com a Nota Final estão neste grupo.

É possível perceber também que dos participantes que não concordaram com a nota final, todos tiveram também algum tipo de discordância (Concordaram Parcialmente ou Não Concordaram) da métrica Pertinência das Mensagens, mais uma vez comprovando o relacionamento entre a aceitação desta métrica com a aceitação da nota final. O Apêndice A apresenta um mapeamento completo das respostas dos participantes sobre cada métrica, para visualização das relações e volume de respostas.

5.2.8 Comparando o Indicador da Participação com as notas atribuídas pelos próprios alunos na autoavaliação

Ao término da sessão os alunos foram convidados a responderem um formulário de autoavaliação da participação na sessão realizada. O objetivo principal foi a possibilidade de comparação das notas atribuídas pelos próprios alunos com as notas geradas pela solução proposta nesta pesquisa e com as notas atribuídas manualmente pelo tutor.

A Tabela 9 apresenta as notas na autoavaliação dos alunos, além de suas respectivas justificativas para as notas enviadas. A nota mais baixa foi atribuída pelo

aluno Alex Brasil (40% de aproveitamento, apenas) que inclusive admite a baixa participação, justificada pelo fato de ter enfrentado problemas técnicos. Através da tabela de mapeamento de respostas apresentada no Apêndice A podemos perceber que apesar de admitir a baixa participação, o participante não concordou com o critério para atribuição da nota final, bem como a nota final atribuída pelo método.

Tabela 9. Notas da autoavaliação

PARTICIPANTE	NOTA	JUSTIFICATIVA
Lucio Camilo	100%	Acredito que eu tenha participado de cada ponto do debate. A turma inteira na verdade participou. Estou me dando 10 mais pelo fato de não saber o que eu não fiz para tirar meus pontos. Repito, acredito ter participado de todos os pontos do debate. É isso.
Mariella	100%	Contribuí com o máximo que eu pude, algumas vezes ficou confuso escrever, pois todos escreviam ao mesmo tempo e não dava tempo de perceber se alguém já tinha citado a informação que eu ainda estava digitando. No geral, acredito que eu tenha visto alguns pontos que passaram despercebidos.
Deise Itália	100%	Considereei que comentei todas as diretrizes que foram apontadas, além de apontar várias. Entrei no horário marcado e me preparei pro debate e acho que fundamentei bem minhas críticas.
Yago Borbato	100%	Participei da conversa, e postei o que era pedido, e cooperei com os demais colegas na tarefa.
Josiel Quintalilha	90%	Eu participei de praticamente todas as fases do bate-papo, estive dentro do assunto, comentei todas as opiniões que foram pedidas e destaquei o que eu acreditava estar errado, a não ser o que já haviam colocado.
Ariel Luis	90%	Acho que mereço uma nota alta por ter lido tudo que foi pedido no livro do Nielsen, portanto me preparando para o bate-papo e para avaliar um site segundo as diretrizes.
Ravi Luis Silva	90%	Creio que minha participação informando minha opinião própria enriqueceram o bate-papo e trouxe novas observações sobre os problemas do site. Tentei debater e dizer em que concordava ou não sobre os problemas mais graves expostos pelos outros. Porém minha falta de alinhamento com as diretrizes lidas não me fez unir os problemas com as próprias diretrizes.
Renata	90%	Acredito que eu tenha ganhado pontos pelas minhas postagens e presença ativa desde o começo do debate, porém eu não acho que respondi em todas as "seções de comentários" - pelo fato de já terem citado o que eu pensava antes, ou ter terminado antes que eu pudesse comentar também - e não disse o nome correto das diretrizes que eu citei. Por isso, então, nota 9.
Daniel Castilho	90%	Por ser uma maneira interativa e de certo modo descontraída, me animei para participar do bate-papo. Achei minha participação muito boa, fiz boa análise do site proposto. Não me daria nota máxima por que sempre há espaço para melhoras, a perfeição é quase impossível.
José Vinícios	90%	Cheguei atrasado, não entendi muito bem mas alcancei. Eu apresentei meus pontos e o que acredito estar correto de acordo com o que fui ensinado nas aulas, e não estava errado, na maior parte. Talvez não tenha sido experiente o bastante pra merecer 10 e devo ter dito algo errado, mas acho que fiz o que devia fazer.

Daivid Angelo	80%	Particpei bem da conversa, apenas tive que procurar o nome das diretrizes em si, pois apenas tinha dado uma breve lida antes.
Geise Fontes	80%	Tentei participar o máximo possível, sem repetir o que os colegas já postaram.
Lucio Salina	80%	Foi uma maneira divertida de aprender, tendo participação do professor e dos alunos, de um modo aonde todos pudessem participar indiferente da quantidade de conhecimento de cada um. Desde a leitura do livro até apontar os erros contribuíram para poder ganhar pontos. Entretanto acho que poderia ter participado ainda mais. Apesar de não ser a melhor plataforma de debates, foi muito produtivo.
Adilson Muller	70%	Eu diria que minha participação no geral foi positiva. Eu realizei uma argumentação baseada no conhecimento proporcionado pelo livro do Nielsen, mas em alguns momentos eu precisaria ter me aprofundado no assunto que estava sendo debatido.
Alex Brasil	40%	Contribuí pouco à discussão, devido a problemas técnicos.

A Tabela 10 compara as notas da autoavaliação com as notas atribuídas pela solução proposta e apresenta a diferença entre as duas notas. Com a tabela, percebemos que 50% da turma teve a diferença entre as duas notas inferior ou igual a 30% e 69% da turma teve a diferença entre as notas inferior ou igual a 40%.

Tabela 10. Comparação entre Autoavaliação e Indicador da Participação

Participante	Indicador da Participação	Nota Autoavaliação	Diferença entre as notas (Indicador da Participação - Autoavaliação)	
Geise Fontes	80%	80%	0%	50% da turma teve diferença entre a autoavaliação e a nota final menor de 30%
Josiel Quintalilha	85%	90%	5%	
Renata	100%	90%	10%	
Daivid Angelo	65%	80%	15%	
Daniel Castilho	65%	90%	25%	
Ravi Luis Silva	65%	90%	25%	
Tarso Nestor	65%	90%	25%	
Adilson Muller	98%	70%	28%	69% teve diferença até 40%
Mariella	65%	100%	35%	
Lucio Salina	46%	80%	35%	
Alex Brasil	4%	40%	36%	Diferença entre as notas maior que 40%. Em alguns casos tivemos quantidade de mensagens baixa, poucas mensagens relevantes e baixa taxa de presença.
Ariel Luis	44%	90%	46%	
Lucio Camilo	41%	100%	59%	
Deise Itália	30%	100%	70%	
José Vinícios	18%	90%	72%	
Yago Borbato	14%	100%	86%	

Todos os participantes que discordaram da nota final atribuída pela solução proposta tiveram a diferença com a autoavaliação maior ou igual a 35% e apenas um

dos participantes que não concordou com a nota final teve a diferença menor que 35% (15% de diferença, com o usuário David Angelo). A partir da tabela acima é possível perceber que a média de distância entre a autoavaliação e a nota estabelecida pelo método de avaliação foi de 36%

No caso onde ocorreu a maior distância entre as notas, participante Yago Borbato, a diferença foi de 86%. Analisando os dados, percebe-se que o participante enviou apenas 10 mensagens durante a sessão e destas, apenas 3 continham palavras consideradas relevantes para a discussão, apesar de ter participado de toda a sessão. O aluno com maior nota atribuída pelo método (100%) enviou 27 mensagens, das quais 17 continham palavras relevantes para o tema. A comparação entre os dois alunos mostra que realmente a autoavaliação de Yago Borbato não foi coerente, considerando a sua participação na sessão.

O aluno José Vinícios (segunda maior diferença – 72%) que se atribuiu 90% na autoavaliação, teve pontuação máxima nos critérios Quantidade de Mensagens e Caracteres, aproveitamento razoável em relação à pertinência das mensagens (65%), mas participou de apenas 28% da sessão. Caso o aluno tivesse participado de toda a sessão, seu aproveitamento nota final seria 65%, o que reduziria a diferença entre a nota final e a autoavaliação para 25%.

A participante Deise Itália teve a nota final igual a 30% e se auto avaliou em 100%. A aluna teve nota máxima em todos os critérios, exceto em relação à pertinência das mensagens, dado que enviou apenas 11 mensagens com palavras relevantes (do total de 37 mensagens enviadas). Analisando manualmente as mensagens enviadas percebe-se que neste caso o método de avaliação proposto não foi justo, pois foram enviadas diversas mensagens que contribuíram para a discussão, mesmo sem mencionar nenhuma das palavras consideradas relevantes pelo tutor.

Os alunos Lucio Camilo e Ariel Luis tiveram situação similar, com 59% e 46% de distância entre as notas, respectivamente. Tais alunos foram penalizados pela solução proposta por terem quantidade de mensagens e/ou caracteres acima da média considerada ideal neste estudo.

O objetivo da autoavaliação foi entender dos próprios alunos como consideraram sua participação na sessão e comparar com a percepção gerada através do método proposto neste estudo. Era esperado que houvesse diferença entre as notas, dado que

normalmente os alunos são otimistas em sua avaliação. No entanto, alguns casos, como o ocorrido com a participante Deise Itália, reforçam a necessidade de aprimoramento dos parâmetros considerados para a qualificação da participação.

5.2.9 Comparando o Indicador da Participação com as notas atribuídas pelo Tutor e Autoavaliação

Após a realização da sessão o tutor foi convidado a avaliar os alunos, atribuindo uma nota individual entre 0% e 100% para a participação na sessão de bate-papo. A avaliação foi realizada de forma livre (manualmente, através da leitura de cada mensagem), para que o tutor considerasse os critérios desejados, sem interferência ou utilização do relatório de avaliação da participação criado nesta pesquisa.

Foi relatado pelo tutor que o tempo total gasto para a avaliação individual dos alunos foi de 1 hora e 13 minutos (mais que o tempo total da sessão) e a nota final foi gerada através de uma média aritmética entre dois fatores:

1) Nota1: Impressão da Participação Individual

Aqui o tutor considerou sua percepção sobre a participação dos alunos durante a sessão, sem considerar critérios específicos ou métricas, apenas considerando sua impressão sobre o desempenho de cada aluno.

2) Nota2: Análise das Postagens

Neste ponto o tutor analisou uma a uma todas as mensagens enviadas pelos alunos para que pudesse concluir sobre a qualidade das mensagens.

A Tabela 11 apresenta as notas dos alunos em cada fator e a média final atribuída pelo tutor.

Tabela 11. Avaliação do Tutor

Nomes	Nota 1	Nota 2	Media Final
Mariella	100%	100%	100%
Deise Itália	95%	95%	95%
Ravi Luis Silva	90%	90%	90%
Renata	90%	90%	90%
Geise Fontes	80%	80%	80%
Lucio Salina	80%	70%	75%
Daivid Angelo	75%	70%	73%

Lucio Camilo	70%	70%	70%
Adilson Muller	65%	65%	65%
Daniel Castilho	65%	65%	65%
Ariel Luis	50%	75%	63%
José Vinícios	80%	29%	55%
Yago Borbato	50%	50%	50%
Josiel Quintalilha	30%	50%	40%
Alex Brasil	5%	5%	5%

A Figura 22 apresenta a distribuição dos alunos em grupos de notas, considerando alunos com nota entre 0% e 35% (nota baixa), 36% a 65% (nota média) e entre 66% e 100% (nota alta). A média da turma de acordo com as avaliações do tutor foi de 68% e a maioria dos alunos (mais da metade) obtiveram bom aproveitamento na avaliação.

■ Entre 66% e 100% ■ Entre 36% e 65% ■ Entre 0% e 35%

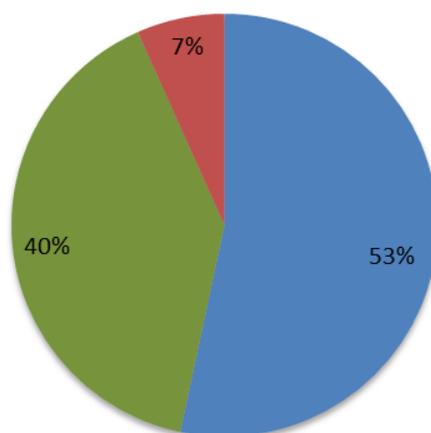


Figura 22. Distribuição das notas dos alunos na visão do Tutor

Na sessão realizada, podemos considerar a nota do tutor como a mais especializada, dado que foi atribuída considerando o que ocorreu durante a sessão, de forma manual, com a análise humana. A comparação das notas da autoavaliação com as notas geradas pela solução proposta nesta pesquisa será realizada nos próximos parágrafos, visando entender a proximidade destas notas.

Ao considerar a média entre os três conjuntos de notas (autoavaliação, nota do tutor e nota calculada pela solução proposta), percebemos que a maior média, corresponde à autoavaliação (o que já era esperado, dado que os alunos normalmente se

avaliam com otimismo). A média de notas da autoavaliação foi de 86%, contra 68% na nota do tutor e 55% na nota obtida através do método proposto nesta pesquisa.

A diferença entre a média das notas do tutor e a média das notas geradas pela solução proposta foi de 13% e a diferença entre a média das notas tutor e autoavaliação foi de 18%. A indicação é que pela média das notas, o método de qualificação da participação proposto nesta pesquisa se aproximou mais da média das notas atribuídas pelo tutor (que é a nota considerada mais acurada).

A Tabela 12 apresenta um quadro com as três notas: Indicador da Participação, Nota do Tutor e Autoavaliação, além de análise de algumas diferenças. Foi feita a análise da diferença entre as notas de cada aluno e, ao final, foram somadas as diferenças para comparação entre as formas de avaliação.

Tabela 12. Comparação entre notas dos alunos (Indicador da Participação, tutor e autoavaliação)

Aluno	Nota Tutor	Indicador da Participação	Auto avaliação	Diferença Indicador da Participação x Tutor	Diferença Auto avaliação x Tutor	Diferença Indicador da Participação x Auto avaliação
Mariella	100%	65%	100%	35%	0%	35%
Deise Itália	95%	30%	100%	65%	5%	70%
Renata	90%	100%	90%	10%	0%	10%
Ravi Luis Silva	90%	65%	90%	25%	0%	25%
Geise Fontes	80%	80%	80%	0%	0%	0%
Lucio Salina	75%	46%	80%	30%	5%	35%
Daivid Angelo	73%	65%	80%	8%	8%	15%
Lucio Camilo	70%	41%	100%	29%	30%	59%
Adilson Muller	65%	98%	70%	33%	5%	28%
Daniel Castilho	65%	65%	90%	0%	25%	25%
Ariel Luis	63%	44%	90%	18%	28%	46%
José Vinícios	55%	18%	90%	37%	36%	72%
Yago Borbato	50%	14%	100%	36%	50%	86%
Josiel Quintalilha	40%	85%	90%	45%	50%	5%
Alex Brasil	5%	4%	40%	1%	35%	36%
Média de Notas	68%	55%	86%	25%	18%	36%

Com a avaliação, tivemos o seguinte cenário:

- Média da Diferença entre Autoavaliação e Notas do Tutor: 18%
- Média da Diferença entre Indicador da Participação e Notas do Tutor: 25%
- Média da Diferença entre Indicador da Participação e Autoavaliação: 36%

Por este cenário, entende-se que a autoavaliação se aproximou mais da nota do tutor que o método proposto nesta pesquisa. Este é mais uma indicação de que é necessário ajustar o método para que as notas se aproximem mais das notas atribuídas pela análise humana do tutor, com olhar mais crítico sobre as mensagens (e capacidade de análise manual mais acurada). A Figura 23 Figura 2apresenta a média da distância entre as notas, com base nas comparações realizadas na Tabela 12.

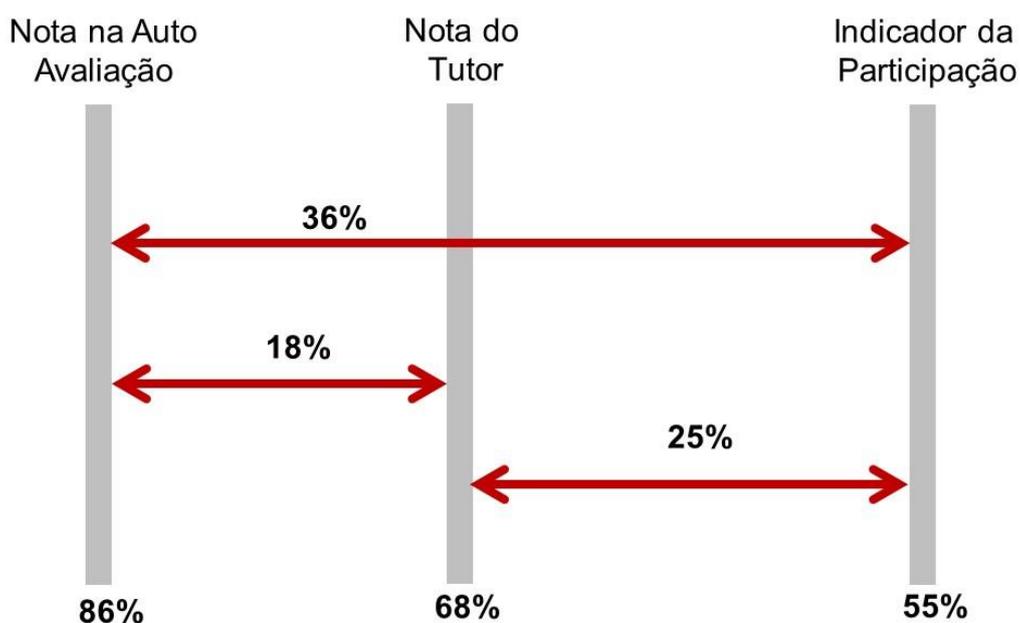


Figura 23. Distância entre a média das notas

Em outra análise, se extrairmos os grupos dos cinco primeiros alunos, com as notas mais altas pelo tutor, e os cinco últimos alunos, com as notas mais baixas também pela avaliação do tutor, percebemos que três dos cinco alunos estão entre os cinco primeiros e cinco últimos alunos, considerando as notas atribuídas com o Indicador da Participação, como está apresentado na Tabela 13. O mesmo ocorre se compararmos o grupo de alunos entre os cinco primeiros e cinco últimos.

Tabela 13. Comparação dos grupos de alunos entre o Indicador da Participação e a avaliação do Tutor

Indicador da Participação		Avaliação do Tutor	
Alunos	Notas	Alunos	Notas
Renata	100%	Mariella	100%
Adilson Muller	98%	Deise Itália	95%
Josiel Quintalilha	85%	Renata	90%
Geise Fontes	80%	Ravi Luis Silva	90%
Mariella	65%	Geise Fontes	80%
Ravi Luis Silva	65%	Lucio Salina	75%

Daivid Angelo	65%	Daivid Angelo	73%
Daniel Castilho	65%	Lucio Camilo	70%
Lucio Salina	46%	Adilson Muller	65%
Ariel Luis	44%	Daniel Castilho	65%
Lucio Camilo	41%	Ariel Luis	63%
Deise Itália	30%	José Vinicios	55%
José Vinicios	18%	Yago Borbato	50%
Yago Borbato	14%	Josiel Quintalilha	40%
Alex Brasil	4%	Alex Brasil	5%

6. Revisão das Métricas e Critérios

Após a aplicação das métricas em estudo de caso real, foi possível identificar algumas possibilidades de melhorias no formato de cálculo do Indicador da Participação. O mesmo ocorreu com algumas métricas e critérios, que não se mostraram adequados e foram criticados pelos participantes. De forma geral, apesar de ter sido percebida proximidade entre o Indicador da Participação e a nota do tutor, ainda existem pontos de melhoria.

Os tópicos a seguir apresentam as melhorias identificadas neste estudo para aumentar a eficiência da solução proposta, aproximando as notas calculadas com a solução proposta nesta pesquisa das notas atribuídas pelo tutor da sessão, consideradas como um tipo de “gabarito” para o estudo.

6.1 Problematizando as métricas propostas

A análise individual das métricas propostas é importante para que sejam identificadas possibilidades de melhorias e ajustes necessários no método de avaliação. Algumas métricas, como Quantidade Média de Caracteres, por exemplo, foram consideradas adequadas para a qualificação da participação. Da mesma forma, a métrica Quantidade de Mensagens (que foi uma escolha óbvia, visto que também foi utilizada por outros estudos similares) teve boa aceitação e não é, aparentemente, um ponto crítico de melhoria.

De acordo com a avaliação das métricas, apresentada no capítulo anterior, os principais pontos de melhoria no método de qualificação da participação proposto estão relacionados com a métrica Pertinência das Mensagens e com o cálculo do Indicador da Participação.

A métrica Pertinência das Mensagens foi a mais polêmica das utilizadas no estudo e foi a responsável direta por algumas das notas baixas atribuídas aos alunos, devido à utilização da estratégia de considerar a presença de palavras consideradas como relevantes pelo tutor nas mensagens enviadas pelos alunos. Esta métrica também

teve um número considerável de alunos que discordaram da nota obtida (13% não concordaram com a nota e 40% concordaram parcialmente).

A fórmula utilizada para o cálculo do Indicador da Participação também gerou dificuldade de entendimento e discordância entre os alunos, sendo a única métrica onde algum participante informou não ter entendido o critério. Além do não entendimento, foi registrado também que alguns participantes não concordaram com o método de avaliação. Esta análise indica a necessidade de melhorar também o método para cálculo da nota final, visando melhorar o entendimento e concordância.

A revisão da métrica Pertinência das Mensagens e da nota final (Indicador da Participação) tem por objetivo aprimorar tais métricas para que as notas geradas pelo método proposto neste estudo sejam mais próximas da avaliação manual do tutor, que está sendo considerada como o “gabarito” da avaliação.

6.2 Melhor definição de palavras chave para a sessão

A definição das palavras chave da sessão levou em consideração um conjunto de palavras previamente definidas pelo tutor, considerando o tema escolhido para a discussão, e a análise dos principais termos utilizados durante a sessão, para refinamento do conjunto inicial de palavras chave. Este conjunto refinado de palavras chave foi o considerado na métrica Pertinência das Mensagens, para o cálculo da quantidade de mensagens com palavras relevantes, com o objetivo de gerar o percentual de pertinência das mensagens de cada participante.

Esta abordagem gerou discordância entre alguns alunos, principalmente entre os que tiveram baixo aproveitamento nesta métrica. Através da análise manual de alguns casos foi identificado que tivemos mensagens relevantes para o tema da sessão que não foram consideradas como pertinentes pela solução proposta, por não possuírem ao menos uma das palavras chave definidas. Este deve ser o principal motivo para as discordâncias relatadas pelos alunos.

Alguns alunos enviaram críticas sobre a métrica, como pode ser visto a seguir:

Aluna Mariella:

Concordo com a minha nota, pois sei que desenvolvi bem o tema, mas obrigar uma pessoa a tratar com palavras pré determinadas

um assunto que pode ser desenvolvido usando sinônimos é um pouco incoerente já que estamos tratando do assunto de forma livre. Para que essa cobrança tivesse coerência o aluno deveria ser notificado de usar um vocabulário técnico antes.

Como se espera que a discussão ocorra de forma livre durante a sessão, a variedade de termos e palavras utilizadas pode ser grande, o que dificulta a qualificação das mensagens por palavras chave. No entanto, caso as palavras chave selecionadas fossem divulgadas os alunos poderiam distorcer a avaliação enviando uma das palavras (mesmo que sem sentido) em cada mensagem.

Deise Itália:

O critério é falho, pois usei diferentes tipos de palavras que não foram consideradas na métrica e fiquei com percentual muito menor que de muitos que não falaram nada novo, somente repetiram o que já havia sido dito. Se eu tivesse mandando todas as mensagens com a palavra site, texto, design, nelas, mesmo não tendo compatibilidade com o tema, minha nota seria maior. Mas usei palavras diferentes e fiquei com nota menor do que quem só mandou algumas poucas mensagens com ideias copiadas ou já ditas ou repetindo a frase anterior.

Este comentário deixa clara a discordância da aluna com a nota recebida. A nota poderia ser melhorada caso o cálculo da nota tivesse se baseado em um conjunto mais refinado de palavras-chave.

Lucio Salina:

As palavras chaves deveriam ser mais abrangentes.

O comentário ratifica a ideia de que é necessário refinar o conjunto de palavras chave para a qualificação da participação, visando tornar a nota dos alunos mais justa nesta métrica. É válido observar que no caso do aluno Lucio Salina o comentário foi enviado mesmo tendo recebido nota máxima no critério.

Outros alunos não discordaram, como ocorreu com Josiel Quintanilha, que informou entender e concordar totalmente com o critério de avaliação e enviou o comentário “Não mudaria nada. O critério está correto.”.

Para o funcionamento adequado desta métrica é necessário definir corretamente o universo de palavras chave para a sessão, aumentando a capacidade de avaliação da pertinência das mensagens enviadas pelos alunos. A liberdade do vocabulário na discussão dificulta a utilização deste tipo de abordagem e aumenta o risco da avaliação, motivo que reforça a necessidade de definição do grupo de palavras chave com cautela e considerando as palavras utilizadas durante a sessão.

Podemos concluir que, com base na revisão das métricas, um conjunto maior de palavras deveria ter sido utilizado para o cálculo do critério de pertinência das mensagens.

6.3 Alteração da métrica Pertinência das Mensagens

A métrica Pertinência das Mensagens considera a quantidade de mensagens enviadas pelos participantes contendo ao menos uma das palavras chave definidas pelo moderador da sessão. Com base na quantidade de mensagens enviadas e na quantidade de mensagens pertinentes, calculamos o percentual de pertinência das mensagens.

Para a atribuição da nota considerou-se que:

- Caso o percentual de pertinência seja igual ou superior a 55%, a nota atribuída para este critério será 100%.
- Caso o percentual de pertinência for igual ou superior a 35% e menor que 55%, a nota atribuída para este critério será 65%.
- Caso o percentual de pertinência for menor que 35%, será considerada nota proporcional variando de 0 a 35% de acordo com a quantidade de mensagens com caracteres relevantes.

Visando melhorar as notas nesta métrica e reduzir a complexidade do cálculo foi avaliada uma alternativa para atribuir 100% à nota caso o aluno tenha enviado mais de 50% de mensagens relevantes e dobrar o percentual de pertinência alcançado, caso seja inferior ou igual a 50%. Neste caso, os resultados seriam:

- Caso o percentual de pertinência seja igual ou superior a 50%, a nota atribuída para este critério será de 100%.

- Caso o percentual de pertinência seja inferior a 50%, a nota atribuída para este critério será o dobro da nota (podendo chegar até 98%, caso a nota original do aluno seja 49%).

A alteração desta métrica teve impacto positivo nas notas de pertinência das mensagens, causando impactos positivos também na nota final. A média de aproveitamento na métrica, que antes do ajuste era de 76%, alcançou 91% após o ajuste, aumentando em 15% a nota média no critério. A nota final foi alterada de 55% para 66% após o ajuste, um aumento de 11%.

A Tabela 14 apresenta os casos com melhora nas notas dos alunos, onde se percebe que o aumento na nota da métrica de pertinência das mensagens e na nota final chegou a 35% do valor original. Dos 15 participantes, 8 (53%) tiveram a nota aumentada após o ajuste no cálculo do critério de Pertinência das Mensagens. A média da diferença das notas no critério de pertinência das mensagens foi de 27% e de 18% na nota final.

Tabela 14. Comparação das notas originais e ajustadas

Alunos	Adequação da Pertinência das Mensagens			Indicador da Participação (Nota Final)		
	Nota Original	Nota Ajustada	Diferença	Nota Original	Nota Ajustada	Diferença
Ravi Luis Silva	65%	100%	35%	65%	100%	35%
Lucio Salina	65%	100%	35%	46%	70%	25%
Alex Brasil	65%	100%	35%	4%	7%	2%
Yago Borbato	30%	60%	30%	14%	28%	14%
Deise Itália	30%	59%	30%	30%	59%	30%
Daivid Angelo	65%	90%	25%	65%	90%	25%
José Vinícios	65%	82%	17%	18%	23%	5%
Mariella	65%	78%	13%	65%	78%	13%

6.4 Alteração do cálculo do Indicador da Participação

O formato aplicado para o cálculo do Indicador da Participação também apresenta necessidade de melhorias, com base no que foi relatado pelos alunos. De acordo com as respostas dos alunos, 5 (33%) entenderam parcialmente e 2 (13%) não entenderam o método aplicado para cálculo da nota final e apenas 54% dos participantes relataram entendimento total do método. A concordância com o critério também foi baixa (28% informaram concordar totalmente) assim como a concordância com a nota

recebida (33% concordaram totalmente e 20% parcialmente, com a discordância alcançando os 47%).

Foram enviados comentários pelos alunos a respeito do cálculo da nota final. Os principais comentários estão descritos a seguir.

Daivid Angelo:

Deveria ser feita uma média e não uma multiplicação de porcentagens.

O aluno obteve 65% na nota final e apresenta comentário sugerindo a utilização da média no cálculo da nota. A utilização de médias aritméticas e ponderadas é comum nas avaliações tradicionais ($\text{Nota 1} + \text{Nota 2} / 2$, por exemplo) e era uma opção lógica que não foi adotada devido ao produto dos percentuais ter maior capacidade de balanceamento das notas (um aluno que tivesse presença de 65% não teria nota superior a este valor mesmo que se tivesse alcançado 100% de aproveitamento em todos os demais critérios).

Alex Brasil:

Alteraria o peso de alguns quesitos, como o de permanência.

Neste caso o aluno sugere a utilização de pesos para a avaliação, remetendo a estratégia da média ponderada. Faz sentido a utilização de média ponderada com peso menor para quesitos como Pertinência das Mensagens, que gerou a maior discordância entre os demais critérios e corresponde a análise qualitativa das mensagens. No entanto, acreditamos que a permanência deve se manter como um critério de peso alto, visto que a presença também é comumente considerada em avaliações tradicionais.

Lucio Salinas:

O modo utilizado para gerar a média não resulta em uma média, fazendo com que pessoas que se destacaram em muitas e teve deslizes em apenas uma tenha uma "média" menor do que o mediano em todas.

Mais uma vez considera-se a utilização da média das notas. A crítica apresentada ataca diretamente a ponderação dos critérios e é comparada com média, que poderia

melhorar a nota de alunos que tiveram bom aproveitamento em alguns critérios e baixo nos demais.

A análise dos comentários enviados pelos alunos indica a possibilidade de substituição do método utilizado para o cálculo da nota final (produto dos percentuais de aproveitamento individuais) por uma média (aritmética ou ponderada). Possivelmente teríamos um melhor entendimento do critério e também melhor concordância com as notas obtidas.

É interessante observar que o formato utilizado pelo moderador da sessão também considera a aplicação de uma média (aritmética, neste caso) para geração da nota final dos alunos. Nota-se que a utilização da média das notas individuais para o cálculo da nota final seria uma escolha simples e de fácil entendimento. Os tópicos a seguir apresentam algumas alternativas estudadas para o cálculo da nota final.

6.4.1 Aplicação de média aritmética para obtenção da nota final

O método desenvolvido para a avaliação participação na sessão considera a avaliação de 4 métricas individuais: Quantidade de Mensagens (Métrica 1), Quantidade Média de Caracteres por mensagem (Métrica 2), Pertinência das Mensagens (Crit. 3) e Tempo de Presença (Crit. 4). A aplicação de média aritmética destes fatores para geração da nota final (Indicador da Participação) foi considerada, gerando a seguinte expressão:

$$\text{Nota Final} = (\text{Crit.1} + \text{Crit.2} + \text{Crit.3} + \text{Crit.4}) / 4$$

A alteração do método para a fórmula acima aumentou a média da turma de 55% para 86%. Caso considerássemos o ajuste feito no critério de Pertinência das Mensagens, a média da turma seria de 89%. O aumento das notas considerando a média aritmética foi considerável (mais de 30%) e acredita-se que aumentaria também a simplicidade do cálculo, melhorando o entendimento e aceitação pelos alunos.

O maior caso de aumento seria para o aluno José Vinícios, que na avaliação original obteve nota final igual a 18% (devido a notas baixas nos critérios individuais) e teria alcançado os 73% na avaliação pela média.

A tabela apresenta o novo quadro de notas, com o ajuste do Indicador da Participação para ser calculado utilizando a média aritmética dos critérios individuais, além da nota original para comparação.

Participantes	% Indicador da Participação - Original	% Indicador da Participação - Média Aritmética	Diferença
José Vinícios	18%	73%	55%
Yago Borbato	14%	68%	54%
Deise Itália	30%	82%	53%
Alex Brasil	4%	51%	47%
Lucio Camilo	41%	83%	42%
Ariel Luis	44%	83%	39%
Lucio Salina	46%	84%	38%
Ravi Luis Silva	65%	91%	26%
Mariella	65%	91%	26%
Daniel Castilho	65%	91%	26%
Daivid Angelo	65%	91%	26%
Geise Fontes	80%	95%	15%
Josiel Quintalilha	85%	96%	11%
Adilson Muller	98%	100%	1%
Renata	100%	100%	0%
Média	55%	86%	31%

6.4.2 Aplicação de média ponderada multiplicada pela presença no bate-papo

Além da aplicação da média aritmética, outra variação considerada foi a média ponderada utilizando os critérios de avaliação definidos. A utilização da média ponderada permite que seja definido um peso menor para o critério de Pertinência das Mensagens, que necessita de refinamento e melhorias conforme discutido anteriormente.

Um fator que consideramos importante na qualificação da participação é o tempo de presença na sessão. Quanto maior o tempo de presença maior será a possibilidade de participar da sessão e contribuir com o aprendizado dos demais participantes. Para aumentar a relevância do tempo de presença na sessão, o retiramos do cálculo da média e multiplicamos o resultado obtido pelo percentual desta métrica individual. Com isso, balancearemos a média pelo tempo de presença.

Os pesos atribuídos para cada métrica foram:

- Quantidade de Mensagens (Crit. 1): Peso 2
- Quantidade Média de Caracteres (Crit. 2): Peso 2
- Pertinência das Mensagens (Crit. 3): Peso 1

A nova fórmula para geração da nota final (Indicador da Participação) seguiu a expressão abaixo:

$$\text{Nota Final} = (((\text{Crit.1*2}) + (\text{Crit.2*2}) + (\text{Crit.3*1})) / 5) * \text{Crit.4})$$

Mais uma vez a alteração do método de cálculo da nota final gerou aumento nas notas dos alunos. A média da turma saltou para 78% (contra os 55% originais) e foi menor que a média apurada pelo método de cálculo pela média aritmética (86%). A média seria ainda maior caso considerássemos o ajuste feito no cálculo da pertinência das mensagens, atingindo os 82%. A diferença média entre a nota final calculada pelo método original e pelo formato utilizando média ponderada e o tempo de presença foi de 24%.

Este formato para cálculo da nota final beneficiou alunos como Deise Itália, que obteve nota máxima em todos os critérios, exceto Pertinência das Mensagens (que obteve 30% de aproveitamento), que foi calculado com peso menor. No caso de Deise Itália, a diferença (que foi a maior) chegou a 56%, saindo de 30% para 86%.

O formato para o cálculo utilizando a média ponderada e a multiplicação pelo tempo de presença parece mais justo que os demais, visto que os pesos podem balancear a fórmula e o tempo de presença pode limitar a nota ao tempo em que o participante esteve presente na sessão.

Tabela 15. Comparativo entre Indicador da Participação

Participantes	% Indicador da Participação - Original	% Indicador da Participação - Média Ponderada	Diferença
Deise Itália	30%	86%	56%
Yago Borbato	14%	63%	49%
Lucio Salina	46%	81%	36%
Lucio Camilo	41%	73%	32%
Ariel Luis	44%	73%	29%

Ravi Luis Silva	65%	93%	28%
Mariella	65%	93%	28%
Daivid Angelo	65%	93%	28%
Daniel Castilho	65%	86%	21%
Alex Brasil	4%	23%	18%
Geise Fontes	80%	92%	12%
Josiel Quintalilha	85%	94%	9%
José Vinicios	18%	26%	8%
Renata	100%	100%	0%
Adilson Muller	98%	98%	0%
Média	55%	78%	24%

7. Conclusão

Nesta pesquisa foram propostas métricas objetivas e critérios para qualificar a participação de alunos em sessões de bate-papo educacional, com o objetivo de apoiar os educadores que utilizam o bate-papo em meio educacional a avaliarem seus alunos. O método de avaliação proposto se baseia na geração de um relatório da participação dos alunos, com o qual educadores podem guiar suas decisões sobre as notas individuais.

A aplicação de estudo de caso exploratório possibilitou a prática do método proposto e entendimento de pontos fortes e fracos, além de possibilidades de melhorias, que foram abordadas nos capítulos anteriores. As seções deste capítulo abordarão os resultados obtidos com a pesquisa e indicações para trabalhos futuros.

7.1 Resultados Obtidos

O método de qualificação da participação elaborado nesta pesquisa baseou-se na avaliação de métricas e critérios individuais com o objetivo de apoiar os educadores na qualificação da participação dos alunos em sessões de bate-papo. Foi gerado um indicador para avaliar a participação, chamado de Indicador da Participação, como resultado da avaliação conjunta das métricas individuais.

A maioria das métricas teve bom nível de entendimento e aceitação pelos alunos participantes. Um formulário aplicado logo após o estudo de caso (bate-papo real com 15 alunos e 1 tutor) mostrou que os alunos concordaram com as notas atribuídas pelos critérios (individuais e nota final) na maioria dos casos.

O estudo de caso mostrou alguns pontos de melhorias (de acordo com as respostas dos formulários pelos participantes), dentre os quais os mais importantes foram:

- Ajuste no cálculo da métrica Pertinência das Mensagens, que pretende avaliar qualitativamente os participantes com base nas mensagens enviadas relevantes enviadas.

- Ajuste no cálculo do Indicador da Participação (nota final), para que se torne mais justo de fácil entendimento pelos alunos.

Foi efetuado ajuste na métrica Pertinência das Mensagens visando simplificar e melhorar o cálculo. Não foi possível realizar um segundo estudo de caso para avaliar novamente as métricas e critérios, mas as alterações que considerando os dados do estudo original, indicam aumento na média de notas do critério (de 15%) e com impacto positivo na nota final (que aumentou 11%).

O método para cálculo do Indicador da Participação (nota final) também foi melhorado, utilizando abordagens diferentes. Foram aplicadas médias aritméticas e ponderadas, considerando critérios como o tempo de presença na sessão como balanceador da nota. A comparação da nota final antes e depois do ajuste indica aumento de mais de 30% na média da turma em um dos casos estudados (quando utilizada média aritmética).

A análise da sessão de bate-papo, que foi o alvo do estudo de caso, também foi feita pelo tutor da sessão, utilizando livre estratégia para avaliação da participação dos alunos. Tal avaliação foi feita pelo tutor em aproximadamente 1 hora e 13 minutos (mais que o tempo da sessão, que foi de 1 hora). Foi visto que o método proposto nesta pesquisa, caso implementado em sistema de computador, poderia gerar o relatório de avaliação em poucos minutos, mesmo considerando atividades manuais do tutor para definição de palavras chave a serem consideradas na análise de pertinência das mensagens.

Comparando a média da turma nas notas do tutor com a média da turma calculada pelo Indicador da Participação (com o ajuste feito no critério de pertinência das mensagens) a diferença é de 4%, menor que a diferença em comparação com a autoavaliação, que é de 6%.

A Fórmula compara as notas atribuídas pelo método proposto nesta pesquisa e seus ajustes com as notas da autoavaliação dos alunos e a avaliação do tutor. Como podemos perceber, a maioria (3 das 4 variações do Indicador da Participação) das variações de notas utilizando o método proposto nesta pesquisa possui mais proximidade da nota do tutor que as notas da autoavaliação.

Tabela 16. Notas gerais e comparações

Participante	(IND 1)	(IND 2)	(IND 3)	(IND 4)	(AA)	(NT)	AA x NT	IND 1 x NT	IND 2 x NT	IND 3 x NT	IND 4 x NT
Adilson Muller	98%	98%	100%	98%	70%	65%	25%	33%	33%	35%	33%
Alex Brasil	4%	7%	51%	23%	40%	5%	35%	1%	2%	46%	18%
Ariel Luis	44%	44%	83%	73%	90%	63%	28%	18%	18%	21%	11%
Josiel Quintalilha	85%	85%	96%	94%	90%	40%	30%	45%	45%	56%	54%
Daivid Angelo	65%	90%	91%	93%	80%	73%	18%	8%	17%	19%	21%
Deise Itália	30%	59%	82%	86%	100%	95%	15%	65%	36%	13%	9%
Mariella	65%	78%	91%	93%	100%	100%	10%	35%	22%	9%	7%
Daniel Castilho	65%	65%	91%	86%	90%	65%	55%	0%	0%	26%	21%
Geise Fontes	80%	80%	95%	92%	80%	80%	70%	0%	0%	15%	12%
José Vinícios	18%	23%	73%	26%	90%	55%	26%	37%	32%	19%	29%
Lucio Camilo	41%	41%	83%	73%	100%	70%	0%	29%	29%	13%	3%
Lucio Salina	46%	70%	84%	81%	80%	75%	5%	30%	5%	9%	6%
Ravi Luis Silva	65%	100%	91%	93%	90%	90%	0%	25%	10%	1%	3%
Renata	100%	100%	100%	100%	90%	90%	0%	10%	10%	10%	10%
Yago Borbato	14%	28%	68%	63%	100%	50%	40%	36%	22%	18%	13%
Média da Turma	55%	64%	85%	78%	86%	68%					
Média de Diferenças							18%	13%	4%	17%	10%

Legenda:

- IND 1: Indicador da Participação (Original)
- IND 2: Indicador da Participação – com ajuste do critério Pertinência das Mensagens
- IND 3: Indicador da Participação – com utilização de Média Aritmética
- IND 4: Indicador da Participação – com utilização de Média Ponderada multiplicada pelo percentual de presença na sessão
- AA: Nota na Autoavaliação
- NT: Nota do Tutor

A Tabela 16 e a **Erro! Fonte de referência não encontrada.** apresentam as médias das diferenças na comparação entre as notas de autoavaliação, avaliações pelo método proposto nesta pesquisa e a nota atribuída pelo tutor. Após a revisão para utilização de média ponderada multiplicada pelo tempo de presença na sessão, foi possível perceber que distância entre o Indicador da Participação e nota atribuída pelo tutor foi reduzida, chegando a 10% na média.

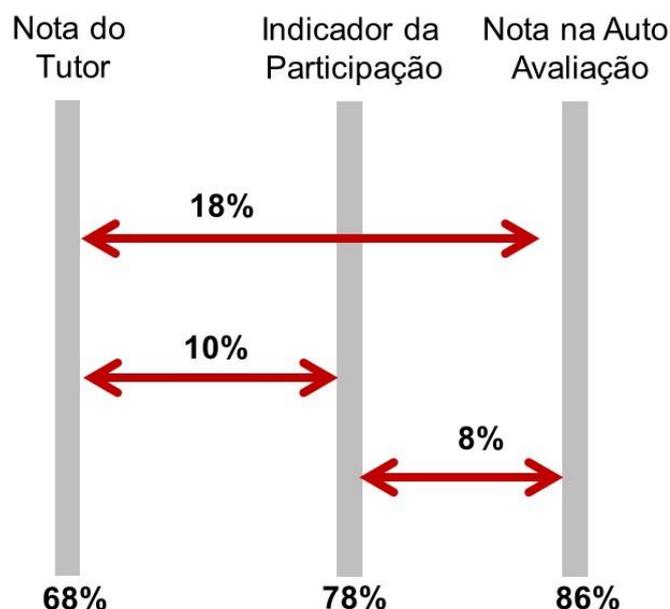


Figura 24. Comparação das diferenças entre as notas

Com isso, percebe-se que o método pode avaliar melhor os alunos que a simples autoavaliação. Se considerarmos que a maioria dos alunos concordou com o método e com as notas obtidas (mesmo no critério inicial, sem ajustes) podemos concluir que as notas calculadas pelo método proposto além de mais próximas das notas do tutor são também aceitas pelos alunos como justas.

7.2 Limitações e Trabalhos Futuros: explorar análise semântica das mensagens

As métricas e critérios estudados estão baseados na qualificação da participação dos alunos tendo como base características quantitativas da sessão, como a quantidade de mensagens e caracteres. A métrica Pertinência das Mensagens foi utilizada no estudo para acrescentar uma análise qualitativa das mensagens enviadas, que poderia ser melhor utilizada na qualificação da participação. Apesar de ter sido a métrica mais criticada, a Pertinência das Mensagens se apresenta como a maior fonte de melhorias para trabalhos futuros.

A análise semântica das mensagens enviadas pelos alunos durante a sessão pode auxiliar a revelar a qualidade das mensagens e se estão realmente relacionadas ao tema em discussão. Com este auxílio, os educadores poderiam utilizar as informações sobre a pertinência das mensagens para visualizar a contribuição dos alunos para a discussão.

Para trabalhos futuros um dos principais pontos a serem explorados é a métrica Pertinência das Mensagens incorporando no processo de geração de informações técnicas de análise semântica das mensagens, mineração de texto, mineração de dados e processamento de linguagem natural. É possível que a qualificação da participação seja feita apenas com esta métrica em alguns cenários, dado que os demais podem ser facilmente burlados por um comportamento padronizado (envio de uma quantidade específica de mensagens e caracteres pode levar o método de qualificação da participação a atribuir uma nota alta para determinado aluno, por exemplo).

7.3 Limitações e Trabalhos Futuros: desenvolvimento de sistema para análise da sessão e geração do Relatório de Performance da Participação

O objetivo principal com a pesquisa foi a definição de métricas objetivas e critérios que permitam a qualificação da participação de alunos em sessões de bate-papo educacional, servindo de apoio aos educadores no momento da avaliação dos alunos. O desenvolvimento das métricas, sem a utilização de um sistema para automatização do processo, pode auxiliar os educadores fornecendo formas de avaliar, mas não resolve o problema relacionado ao tempo necessário para a avaliação.

A métrica Pertinência das Mensagens envolveu a análise das palavras enviadas durante a sessão e processos de mineração de texto para aprimoramento do conjunto de palavras mais utilizadas na sessão. Algumas destas palavras foram utilizadas para refinar o grupo de palavras chave consideradas, as incluindo na lista inicial definida pelo moderador. Para este processo foi desenvolvida rotina de avaliação da sessão para cálculo do percentual de mensagens relevantes por participante.

Não foi possível desenvolver durante a pesquisa um sistema para automatização de todo o processo, desde a análise das mensagens até a geração do relatório de performance da participação. Com tal limitação, diretamente ligada ao tempo necessário para construção do sistema, os cálculos para geração das notas em cada métrica foram criados em planilha eletrônica (exceto para a métrica Pertinência das Mensagens, que teve apoio de rotina de programação) tanto para as notas individuais em cada métrica quanto para a nota final. Após o cálculo das notas, o relatório de avaliação foi gerado manualmente e enviado aos alunos.

A implementação das regras para cálculo das notas e geração do relatório de performance da participação em sistema disponível para o tutor tornaria o processo mais completo. O tutor poderia informar o grupo de palavras chave e todos os cálculos seriam realizados automaticamente pelo sistema (visto que apenas este critério depende de interação com usuário), além de gerar o relatório da participação. Com isso, os usuários poderiam inclusive visualizar no sistema os gráficos de avaliação e descobrir suas notas, tornando o processo mais interativo.

As métricas definidas poderiam ter critérios editados pelo tutor durante a avaliação da sessão, ajustando, por exemplo, os limites de mensagens e caracteres de acordo com o entendimento da sessão, além da definição das palavras chave. Os horários de início e término da sessão poderiam também ser ajustados para remover períodos de boas vindas e encerramento, por exemplo.

Sendo assim, uma das propostas para trabalhos futuros consiste no desenvolvimento do sistema para análise dos dados e geração automática do relatório de performance da participação, com a disponibilização de tal sistema aos educadores via Internet, de forma que agregue valor à utilização de sessões de bate-papo em meio educacional.

7.4 Limitações e Trabalhos Futuros: execução de outros estudos de caso para avaliar as métricas ajustadas

Durante a fase inicial da pesquisa foram definidos métricas e critérios que seriam utilizados para a qualificação da participação dos alunos na sessão de bate-papo. Tais métricas foram definidas considerando estudos anteriores para o entendimento de quais seriam os melhores critérios a serem aplicados e quais limites poderiam ser considerados. A aplicação do estudo de caso descrito neste trabalho foi essencial para a validação das métricas criadas e do processo de qualificação da participação.

Após o estudo de caso foram analisadas as respostas dos alunos sobre cada critério de avaliação e os resultados obtidos, com o objetivo de melhorar as métricas e os critérios considerados. A revisão destas métricas permitiu maior proximidade entre as notas atribuídas pela solução proposta nesta pesquisa e as notas definidas através da avaliação humana do tutor. As métricas e critérios ajustados compõem um conjunto

mais confiável de regras que podem ser utilizadas para qualificar a participação dos alunos em sessões de bate-papo educacional.

A execução de novo estudo de caso tendo como base as métricas já ajustadas poderia reforçar a aplicabilidade do método desenvolvido e apontar novas oportunidades de refinamento. Outros pontos de vista também poderiam ser analisados, como o nível de aceitação das métricas pelo tutor da sessão e as possíveis melhorias na abordagem criada.

Um estudo ainda maior poderia ser conduzido com a execução de diversas sessões de bate-papo, para assuntos e públicos variados visando gerar maior volume de informações sobre o entendimento e aceitação das métricas, bem como aceitação das notas atribuídas pelo método de qualificação da participação proposto. Idealmente os novos estudos poderiam ser conduzidos utilizando sistema para a automatização da geração do relatório da participação, de forma que o processo geral pudesse ser avaliado.

O segundo estudo de caso (e os possíveis subsequentes) não foi executado dentro da pesquisa por limitações de prazo, visto o tempo necessário para preparação das sessões, análise imediatamente posterior dos dados para a geração do relatório de performance da participação, análise das opiniões dos alunos e resultados obtidos.

Como sugestão para trabalhos futuros, além do sistema para automatizar a geração do relatório da participação, é proposta a realização de outras sessões reais de bate-papo educacional. Com maior volume de dados derivado de novas sessões seria possível gerar mais evidências da relevância desta pesquisa e identificação de pontos de melhoria para a criação de produto que auxilie os educadores em seu dia a dia na utilização de sessões de bate-papo educacional.

Referências

- ACESSOS. Serviço disponibilizado no TelEduc para verificar os acessos e a frequência dos alunos e formadores no ambiente virtual de aprendizado. Disponível em <http://penta3.ufrgs.br/tutoriais/teleducv3/acessos.htm>. Último acesso: 17/10/2014.
- ADAMS, P. H. & MARTELL, C. H. (2008) Topic Detection and Extraction in Chat. II IEEE International Conference on Semantic Computing. Santa Calara, CA – USA.
- ANJEWIERDEN, A., KOLLOFFEL, B., & HULSHOF, C. (2007). Towards educational data mining: Using data mining methods for automated chat analysis to understand and support inquiry learning processes. In International Workshop on Applying Data Mining in e-Learning (ADML 2007).
- BARILLI, E. C. V. C. “Avaliação: acima de tudo uma questão de opção”. In: SILVA e SANTOS (orgs), *Avaliação da Aprendizagem em Educação online*. São Paulo: Loyola, 2006.
- CALDEIRA, A. C. M. (2004). Avaliação da aprendizagem em meios digitais: novos contextos. <http://www.abed.org.br/congresso2004/por/htm/033-TC-A4.htm>. Último Acesso em: 11/08/2014.
- DA ROCHA, H. V., OTSUKA, J. L., FREITAS, C. E., & FERREIRA, T. B. “Avaliação online: o modelo de suporte tecnológico do projeto telEduc”. In: SILVA e SANTOS (orgs), *Avaliação da Aprendizagem em Educação online*. São Paulo: Loyola, 2006.
- ELLIS, C.A., GIBBS, S.J. & REIN, G.L. Groupware - Some Issues and Experiences. In: Communications of the ACM, v. 34, n. 1. 1991, p. 38-58.
- ESTRUC, M., PIMENTEL, M. (2012). Portal Tagarelas: bate-papo para educação. XXIII Simpósio Brasileiro de Informática na Educação – SBIE. Rio de Janeiro, RJ.
- FERNANDES, D. (2006). Para uma teoria da avaliação formativa. Revista Portuguesa de Educação, v. 19, n. 2, p. 21-50.
- FUKS, H. (2000). Aprendizagem e trabalho cooperativo no ambiente AulaNet. Revista Brasileira de Informática na Educação, 6(1), 53-74.

- FUKS, H., PIMENTEL, M. G., GEROSA, M. A., FERNANDES, M. C. P., & DE LUCENA, C. J. P. “Novas Estratégias de Avaliação Online: aplicações e implicações em um curso totalmente a distância através do ambiente AulaNet”. In: SILVA e SANTOS (orgs), *Avaliação da Aprendizagem em Educação online*. São Paulo: Loyola, 2006.
- GAINARU, A. *et al.*(2010). Toolkit for automatic analysis of chat conversations. VIII International Conference on Communications – COMM. Bucharest, Romania.
- HRASTINSKI, S. (2008). What is online learner participation? A literature review. *Computers & Education*
- HUI, S. C., HE, Y., & DONG, H. (2008). Text mining for chat message analysis. In *Cybernetics and Intelligent Systems, 2008 IEEE Conference on* (pp. 411-416). IEEE.
- INTERMAP. Serviço disponibilizado no TelEduc para visualização da interação dos participantes de cursos on-line. Disponível em <http://penta3.ufrgs.br/tutoriais/teleducv3/intermap.htm>. Último acesso: 17/10/2014.
- OH, S., LICHTNOW, D., SALDAÑA, R., BORGES, T., PRIMO, T., KICKHÖFEL, R., SIMÕES, G. (2004) Investigação sobre a Identificação de Assuntos em Mensagens de Chat. XXIV Congresso da Sociedade Brasileira de Computação. Salvador, BA.
- LUCKESI, C. C. (2011). Avaliação da aprendizagem: componente do ato pedagógico. Cortex.
- MOODLE. Página sobre utilização do componente de bate-papo. Disponível em http://docs.moodle.org/27/en/Using_Chat. Último acesso: 26/07/2014.
- NETMINER. Software para análise de redes sociais. Disponível em <http://www.netminer.com/index.php>. Último acesso: 31/07/2014.
- NUNES, L. C., VILARINHO, L. R. G. “Avaliação da aprendizagem no ensino online em busca de novas práticas”. In: SILVA e SANTOS (orgs), *Avaliação da Aprendizagem em Educação online*. São Paulo: Loyola, 2006.
- PERRENOUD, P. (1999). Avaliação: da excelência à regulação das aprendizagens-entre duas lógicas. Artmed.
- PERRENOUD, P., THURLER, M. G. (2002). As competências para ensinar no século XXI: a formação dos professores e o desafio da avaliação. Artmed.

- PIMENTEL, M. G., FUKS, H., & LUCENA, C. J. P. D. (2003a). “Debati, debati... aprendi? Investigações sobre o papel educacional das ferramentas de bate-papo”. In Anais do Workshop de Informática na Escola (Vol. 1, No. 1, pp. 61-72).
- PIMENTEL, M. G., FUKS, H., & LUCENA, C. J. P. D. (2003b). Avaliação da Participação dos Aprendizes em Debates Síncronos. In Anais do Simpósio Brasileiro de Informática na Educação (Vol. 1, No. 1, pp. 136-145).
- PIMENTEL, M., FUKS, H., & LUCENA, C. J. P. (2004). Avaliação da Participação em Conferências Textuais Assíncronas. In Anais Eletrônico do X Workshop de Informática na Escola, integrante do XXIV Congresso da Sociedade Brasileira de Computação (WIE/SBC 2004), ISBN (pp. 85-88442).
- RABBANY, R., TAKAFFOLI, M., & ZAIËANE, O. R. (2011). Analyzing participation of students in online courses using social network analysis techniques. In Proceedings of educational data mining.
- ROCHA, E.B. (2013) Modelo de Participação em Bate-Papo Educacional. Dissertação de Mestrado. Departamento de Informática Aplicada, UNIRIO.
- SANTOS, L. (2002). Autoavaliação regulada: porquê, o quê e como?
- SANTOS, E. (2010). Educação online para além da EAD: um fenômeno da cibercultura. SILVA, M., PESCE, L.; ZUIN, A. Educação Online: cenário, formação e questões didático metodológicas. Rio de Janeiro: Wak.
- SILVA, D. R., SENO, W. P., VIEIRA, MARINA, TP. (2001). “Acompanhamento do Aprendizado em Educação a Distância com Uso de Data Mining.”. In: Conferência Latinoamericana de Informática, Mérida, Venezuela. 2001.
- SILVA, M. (2006). O fundamento comunicacional da avaliação da aprendizagem na sala de aula online. *Avaliação da aprendizagem em educação online. São Paulo: Loyola.*
- TELEDUC. Ambiente para realização de cursos on-line. Disponível em http://ggte.unicamp.br/~teleduc/pagina_inicial/teleduc.php. Último acesso: 17/10/2014.
- UCINET. Software para análise de redes sociais. Disponível em <https://sites.google.com/site/ucinetsoftware/home>. Último acesso: 31/07/2014.

WORDNET. Disponível em <http://wordnet.princeton.edu>. Último acesso: 26/07/2014.

ZAIANE, O. R., & LUO, J. (2001). "Towards evaluating learners' behaviour in a web-based distance learning environment. In *Advanced Learning Technologies*", 2001. Proceedings. IEEE International Conference on (pp. 357-360). IEEE.

Apêndice A – Mapeamento de respostas dos alunos sobre a avaliação das métricas

A Tabela 17 Tabela 1abaixo apresenta o mapeamento de todas as respostas dos alunos para cada métrica de qualificação da participação, de acordo com os questionamentos do formulário de avaliação do relatório de Performance da Participação, aplicado após a sessão de bate-papo realizada no estudo de caso.

Tabela 17. Mapeamento de Respostas

Participante	Concorda com a métrica Quantidade de Mensagens?	Concorda com sua nota na Quantidade de Mensagens?	Concorda com a métrica Quantidade Média de Caracteres?	Concorda com sua nota na Quantidade Média de Caracteres?	Concorda com a métrica Pertinência das Mensagens?	Concorda com sua nota na Pertinência das Mensagens	Concorda com o critério Tempo de Presença na Sessão?	Concorda com sua nota no Tempo de Presença na Sessão?	Concorda com o critério para a Nota Final?	Concorda com sua Nota Final (Performance da Participação)?
Josiel Quintalilha	Concordo parcialmente	Concordo parcialmente	Concordo parcialmente	Concorda totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente
Geise Fontes	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concorda totalmente	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente
Mariella	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concorda totalmente	Concordo parcialmente	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Não concordo	Não concordo
Renata	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concorda totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concordo totalmente
Deise Itália	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Não concordo	Não concordo	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Não concordo
Yago Borbato	Concordo totalmente	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Não concordo
Daivid Angelo	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Não concordo	Não concordo
Adilson Muller	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente
José Vinícios	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concordo parcialmente	Concordo parcialmente	Concordo parcialmente	Não concordo	Não concordo
Alex Brasil	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Não concordo	Concordo parcialmente	Concordo totalmente	Concordo parcialmente	Não concordo	Não concordo
Lucio Salina	Concordo parcialmente	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Não concordo	Não concordo
Lucio Camilo	Concordo parcialmente	Concordo parcialmente	Concordo totalmente	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concordo parcialmente
Ravi Luis Silva	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concordo parcialmente
Daniel Castilho	Concordo parcialmente	Não concordo	Concordo parcialmente	Concordo totalmente	Concordo parcialmente	Concordo totalmente	Concordo totalmente	Concordo totalmente	Concordo parcialmente	Concordo parcialmente
Ariel Luis	Não concordo	Não concordo	Não concordo	Não concordo	Não concordo	Não concordo	Não concordo	Não concordo		Concordo totalmente

Apêndice B – Relatório de Performance da Participação

Durante a pesquisa foi discutido o relatório gerado como resultado da qualificação da participação dos alunos, visando apresentar o panorama da participação dos alunos em cada métrica estabelecida e apoiar, assim, o educador na atribuição de nota aos alunos. O relatório aplicado no estudo de caso foi reproduzido abaixo, para conhecimento, removendo apenas os nomes dos participantes em cada gráfico.

Relatório de Performance da Participação no bate-papo

O relatório a seguir contém a avaliação da **participação** dos envolvidos no bate-papo. Ele está separado por áreas, onde serão apresentadas as avaliações parciais e ao final é apresentada a nota final da Performance da Participação na sessão.

A cada área será apresentado um breve resumo dos critérios utilizados na avaliação, para facilitar o entendimento.

Lembre-se, na avaliação do relatório e no preenchimento do questionário, que o objetivo é a avaliação da Participação na sessão, bem como sua contribuição e não se as mensagens estavam corretas ou erradas para as tarefas dadas pelo moderador.

1. Critério 1: Adequação à Quantidade de Mensagens enviadas

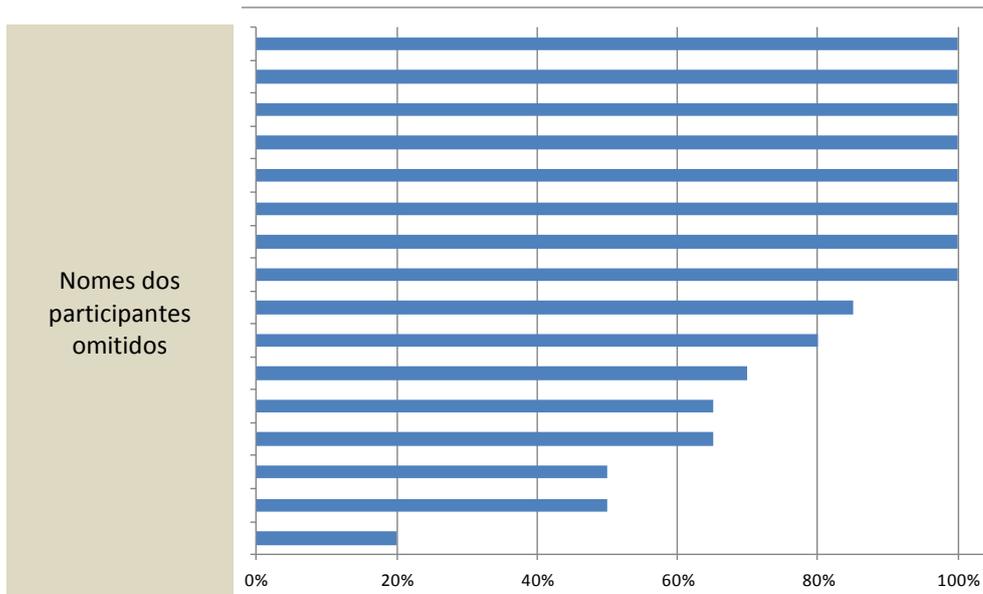


Gráfico 1

Métricas utilizadas:

A avaliação da quantidade de mensagens foi realizada considerando a proximidade à quantidade de mensagens entendida como ideal na sessão. Foram utilizados os critérios:

1. De 0 a 19 mensagens: Nota proporcional (de 0 a 99%, de acordo com a quantidade de mensagens enviadas, onde para nenhuma mensagem a nota será próxima de 0% e para 19 mensagens a nota será próxima de 99%);
2. De 20 a 50 mensagens: Nota 100% (em estudos anteriores vimos que a média de mensagens por participante em sessões é de aproximadamente 35 mensagens por sessão);
3. De 51 a 100 mensagens: Nota proporcional (de 0 a 99%, de acordo com a quantidade de mensagens enviadas, onde para 100 mensagens a nota será próxima de 0% e para 51 mensagens a nota será próxima de 99%);
4. Acima de 100 mensagens (mais que o dobro do limite considerado com o bom): Nota 0%, visto que excede a quantidade de mensagens aceitável dentro da sessão;

2. Critério 2: Adequação à Quantidade de Caracteres por Mensagem

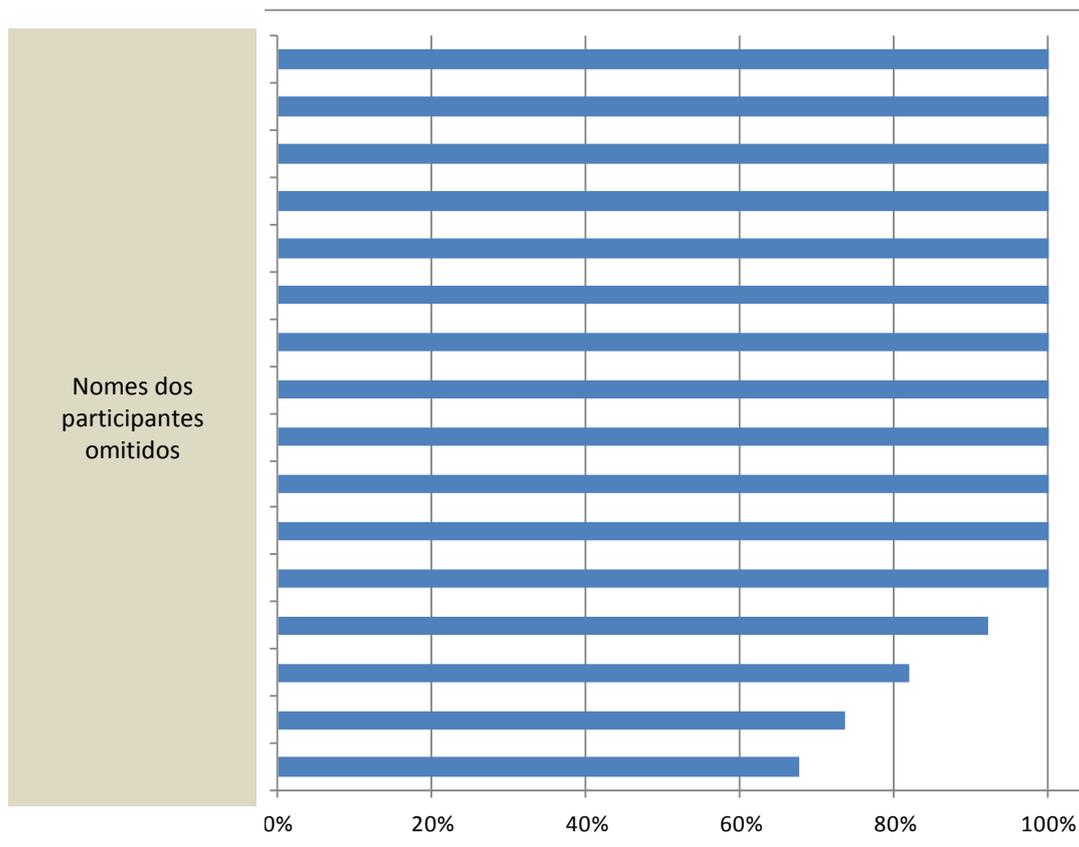


Gráfico 2

Métricas utilizadas:

A avaliação da adequação da quantidade média de caracteres enviados foi realizada considerando a proximidade da quantidade de caracteres média por mensagem entendida como ideal na sessão. Foram utilizados os critérios:

5. De 0 a 34 caracteres por mensagem: Nota proporcional (de 0% a 99%, de acordo com a quantidade de caracteres enviados, onde para nenhum caractere a nota será próxima de 0% e para 34 caracteres em média por mensagem a nota será próxima de 99%);
6. De 35 a 105 caracteres por mensagem: Nota 100% (quantidade considera boa para o tamanho das mensagens);
7. De 106 a 210 caracteres por mensagem: Nota proporcional (de 0% a 99%, de acordo com a quantidade de mensagens enviadas, onde para 210 caracteres em média por mensagem a nota será próxima de 0% e para 106 caracteres em média por mensagem a nota será próxima de 99%);
8. Acima de 210 caracteres em média por mensagens (mais que o dobro do limite considerado com o bom): Nota 0%, visto que excede a quantidade de caracteres média por mensagem aceitável dentro da sessão;

3. Critério 3: Percentual de Pertinência das Mensagens

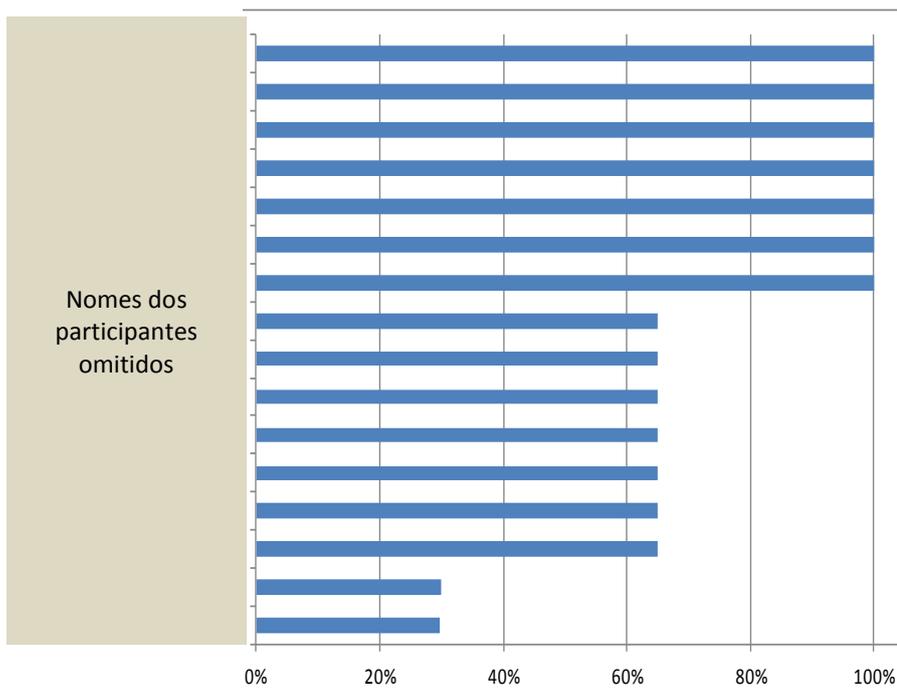


Gráfico 3

Métricas utilizadas:

A avaliação no critério Pertinência das Mensagens foi realizada considerando o percentual de mensagens enviadas que contém ao menos uma das palavras consideradas como relevantes (vide quadro abaixo) para o tema em discussão (quantidade de mensagens relevantes / quantidade total de mensagens enviadas pelo participante).

A partir deste percentual, foram considerados os seguintes critérios:

- Caso o percentual de pertinência for $\geq 55\%$ → Nota utilizada para avaliação final: 100%
- Caso o percentual de pertinência for $\geq 35\%$ → Nota utilizada para avaliação final: 65%
- Caso o percentual de pertinência for $< 35\%$ → Nota proporcional para avaliação final (de 0 a 35%)

As palavras consideradas como importantes para o tema foram:

link,links,site,sites,cor,texto,textos,usuário,usuários,usuarios,visual,pagina,problema,diretriz ,diretrizes,elementos, fonte,design,conteúdo,conteudo,erros,letra,navegador,solução,solucao,usabilidade

4. Critério 4: Percentual de Permanência na Sessão

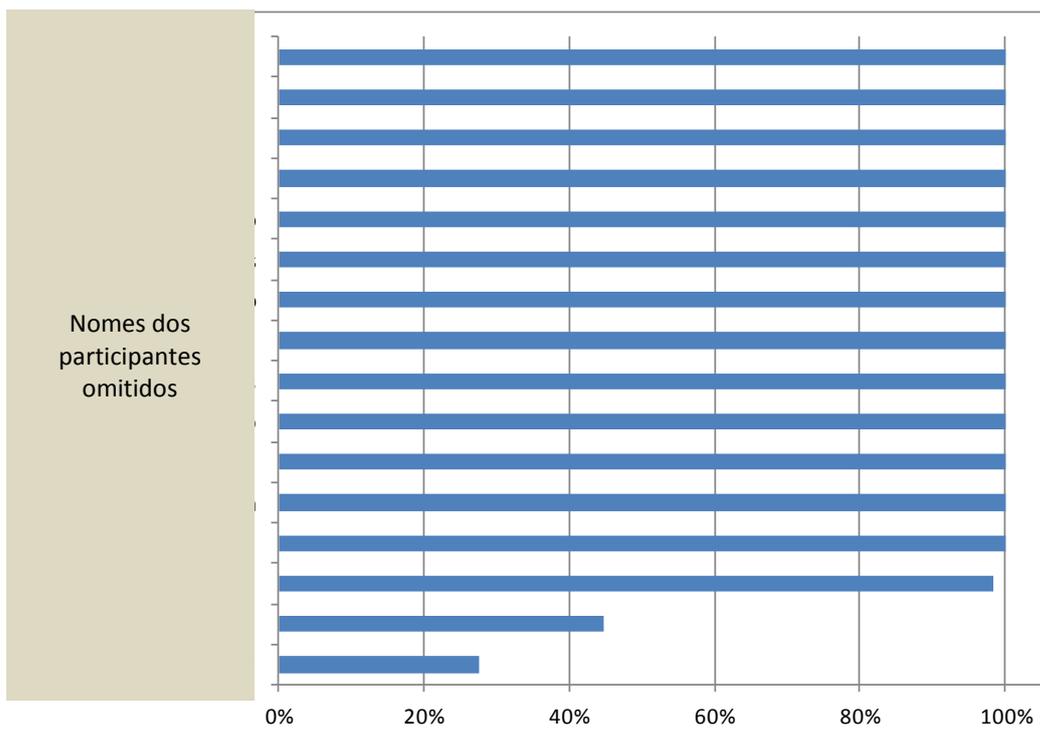


Gráfico 4

Métricas utilizadas:

A avaliação no critério Percentual de Permanência na Sessão das Mensagens foi realizada considerando o percentual do tempo de participação do usuário.

O tempo de permanência foi calculado com tendo em vista o momento da primeira mensagem enviada pelo participante na sessão e o horário de término da sessão (horário da primeira mensagem – horário de término da sessão).

Participantes que estavam presentes (responderam a chamada inicial) antes do início formal da sessão foram contabilizados como presentes desde o início.

5. NOTA FINAL - Indicador da Participação

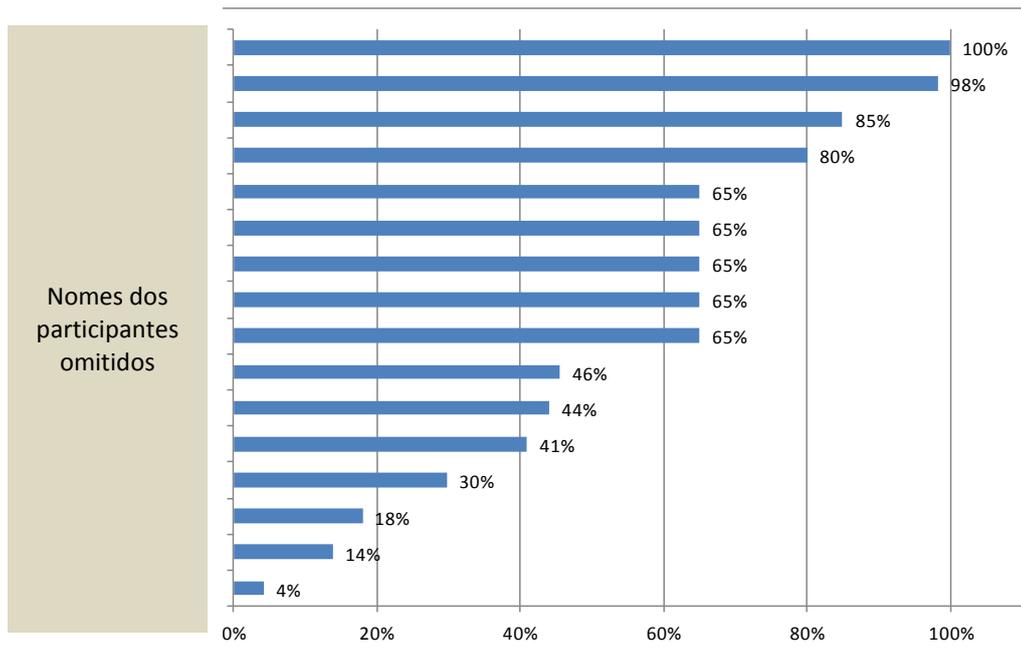


Gráfico 5

Métricas utilizadas:

A nota final na avaliação da participação na sessão corresponde à Performance da Participação, que foi calculada com a seguinte fórmula:

$$\text{Critério 1} \times \text{Critério 2} \times \text{Critério 3} \times \text{Critério 4}$$

Assim, o percentual é calculado com base em todos os critérios, refinando a avaliação. Caso o participante tenha obtido 100% no critério "Adequação à Quantidade de Mensagens enviadas" a nota final será ajustada de acordo com o percentual de aproveitamento dos outros critérios.

6. Notas no Formulário de Autoavaliação

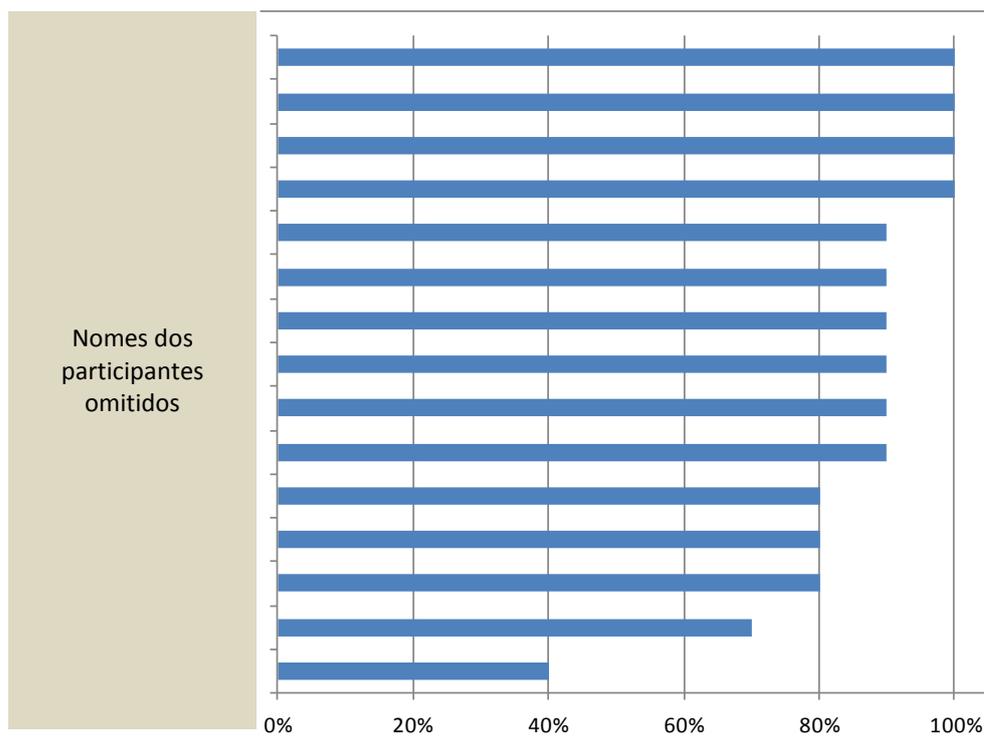


Gráfico 6

OBS.: As notas acima correspondem às notas informadas no Formulário de Autoavaliação preenchido após a sessão.