

Design Science Research para o desenvolvimento de um Modelo da Participação em Bate-papo

Edmilson Barcelos Rocha, Mariano Pimentel, Morganna Carmem Diniz, Flávia Maria Santoro

Departamento de Informática Aplicada – Universidade Federal do Estado do Rio de Janeiro (UNIRIO) – Rio de Janeiro – RJ – Brasil

{edmilson.rocha,pimentel,morganna,flavia.santoro}@uniriotec.br

Resumo. *O grande crescimento da Educação a Distância no Brasil e o intenso uso do bate-papo nessa modalidade motivaram o estudo apresentado neste artigo. A pesquisa tem um duplo objetivo: demonstrar a influência do tamanho do grupo na participação em bate-papo educacional e produzir um modelo matemático para estimar a quantidade máxima de alunos que devem participar de um bate-papo mantendo o nível de participação desejado pelo professor. A metodologia usada foi a Design Science Research, que orienta a realização de uma pesquisa comportamental aliada à produção de um artefato. As estimativas produzidas pelo modelo foram confrontadas com dados reais e mostraram precisão adequada corroborando para sua validade.*

Abstract. *The growth of distance education in Brazil and the intensive use of chat in this context motivated the study presented in this paper. The research has two objectives: investigate the influence of group size on participation in chat and produce a mathematical model to estimate the maximum number of students who must participate in a chat session but maintaining the desired level of participation set by the teacher. The methodology used was Design Science Research, which combines behavioral research with the production of an artifact. The results produced by the model were compared with real data and showed adequate accuracy, which confirms the validity of the model.*

1. Introdução

A motivação para a realização da pesquisa apresentada neste artigo é gerar conhecimento sobre o uso do bate-papo na educação. A importância dessa pesquisa decorre do grande crescimento da Educação a Distância no Brasil e o intenso uso de bate-papo nessa modalidade.

A partir do ano 2000 ocorreu um grande crescimento da EAD no Brasil: de praticamente zero cursos a distância, em um pouco mais de uma década esta modalidade passou a ser responsável por 15,8% das matrículas na graduação, ultrapassando um milhão de estudantes [INEP 2013]. A mídia *online* já é o recurso educacional mais utilizado pelas instituições que oferecem cursos a distância [Censo EAD.BR, 2013]. Os recursos *online* têm ganhado cada vez mais espaço por favorecerem um novo paradigma de educação mais voltado para a interação. O bate-papo é o terceiro meio de interação *online* mais utilizado na EAD, atrás apenas do fórum e do e-mail conforme dados apresentados no CENSOEAD.BR (2013, p.163).

Na educação *online*, o uso de sistemas de comunicação síncrono, especialmente

o bate-papo, é recomendado por pesquisadores por promover o aumento da participação dos alunos [Hrastinski 2007; Stahl 2009]. Mas quantos alunos devem participar de um bate-papo realizado no contexto educacional? Saber se numa sessão de bate-papo podem participar ao mesmo tempo 100 ou apenas 10 alunos tem um grande impacto financeiro e no planejamento das atividades de um curso a distância. A informação do limite de participantes em uma sessão de bate-papo educacional é encontrada na literatura de forma imprecisa e conflitante, o que representa um problema relevante, discutido na Seção 3 do presente artigo.

Para saber quantos alunos devem participar de uma sessão de bate-papo educacional, foi construído um modelo matemático, apresentado na Seção 4. O modelo concebido, baseado na Teoria das Filas, é um artefato capaz de produzir estimativas objetivadas e corroborar com a pesquisa comportamental sobre a influência do tamanho do grupo na participação em bate-papo educacional. O modelo foi desenvolvido por meio de um processo iterativo caracterizado pelo ciclo construção-avaliação, conforme abordado na Seção 5. A avaliação da precisão das estimativas produzidas pelo modelo é apresentada na Seção 6. Na seção 7 são discutidas as contribuições desta pesquisa realizada, que visa tanto o público técnico de modelagem de sistemas computacionais e de comunicação quanto o público gerencial de planejamento de cursos de educação a distância.

A presente pesquisa foi conduzida seguindo as diretrizes de Design Science Research, que é uma abordagem tipicamente adotada na área de Sistemas de Informação, conforme discutido na próxima seção.

2. Design Science Research

A pesquisa foi guiada pelas diretrizes do método Design Science Research, proposto na área de Sistemas de Informação [Hevner et al. 2004]. De acordo com Hevner e colaboradores, os conhecimentos necessários para realizar uma pesquisa em SI envolvem dois paradigmas complementares: Ciência do Comportamento e Ciência do Design. O paradigma da ciência do comportamento aborda a pesquisa pelo desenvolvimento de teorias que explicam ou predizem fenômenos relacionados à necessidade de negócio identificada. Já o paradigma da ciência de design aborda a pesquisa pelo desenvolvimento e avaliação de artefatos projetados para atender à necessidade de negócio identificada.

Na presente pesquisa, o fenômeno comportamental investigado é a influência do tamanho do grupo na participação em bate-papo educacional; e o artefato desenvolvido é um modelo matemático da participação nessas sessões de bate-papo. Em Design Science Research, o artefato é projetado com base nas conjecturas teóricas fundamentadas na Ciência do Comportamento; e a avaliação do artefato fornece novos dados sobre as conjecturas elaboradas visando aumentar o conhecimento teórico. Na Figura 1, é ilustrado o ciclo Design Science Research adaptado à presente pesquisa.

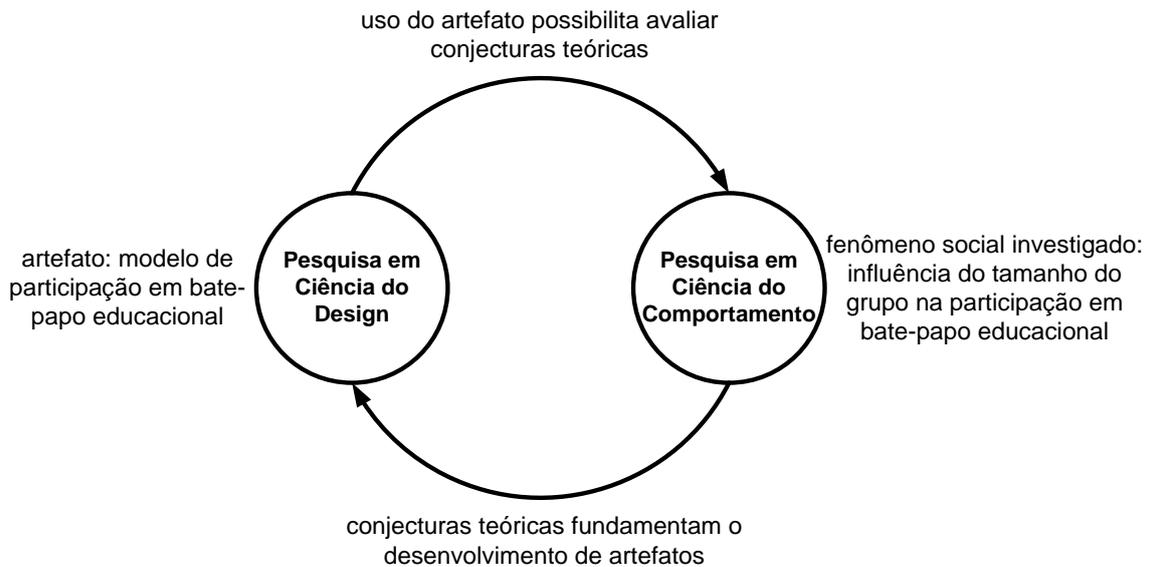


Figura 1. Ciclo DSR adaptado à pesquisa

A Design Science Research busca elevar o desempenho de pesquisas em Sistemas de Informação por meio de um framework conceitual conciso para entender, executar e avaliar a pesquisa. O framework conceitual apresentado por Hevner e colaboradores (2004) e adaptado a presente pesquisa conforme Figura 2, combina Ciência Comportamental e Ciência de Design. De acordo com o framework, uma pesquisa é realizada em um ambiente produzindo uma base de conhecimento. Os produtos da pesquisa são: Teorias e Artefatos que devem ser Justificados e Avaliados para produzirem *feedback* que possam refinar as Teorias e os Artefatos.

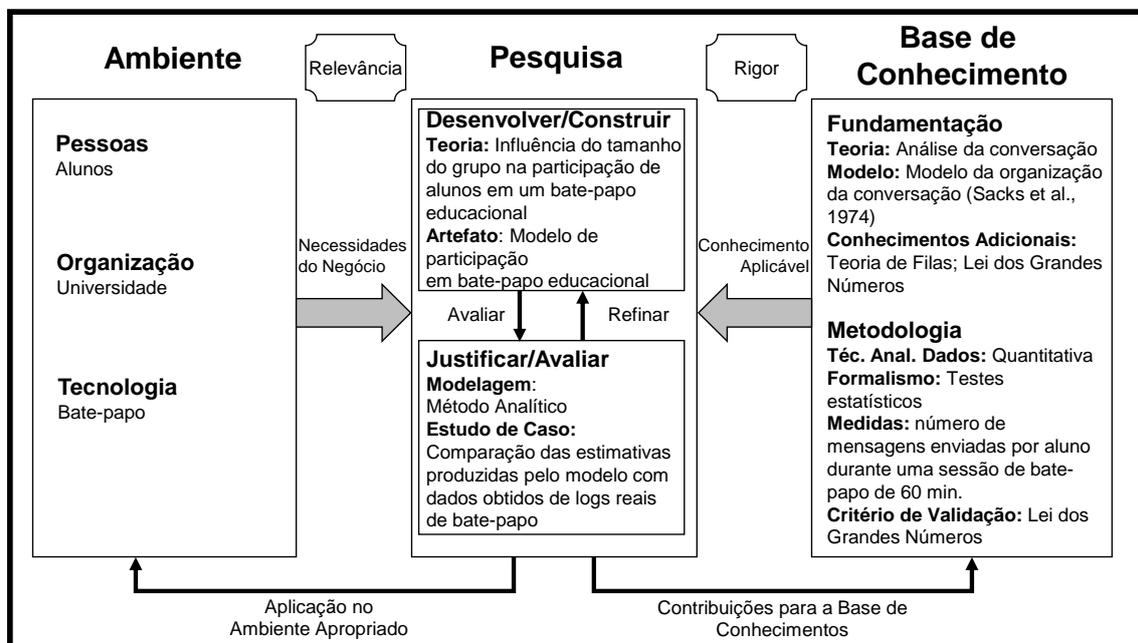


Figura 2. Framework da DSR instanciado para a presente pesquisa

Design Science Research também disponibiliza um conjunto de diretrizes para as pesquisas em Sistema de Informação. Na Tabela 1, são listadas as diretrizes propostas por Hevner e colaboradores (2004) e a aplicação dessas diretrizes na presente pesquisa.

Tabela 1. Diretrizes do Design Science Research

DIRETRIZ	APLICAÇÃO DA DIRETRIZ NESSA PESQUISA
Relevância do Problema (Seção 3)	Problema: Não se sabe ao certo quantas pessoas devem participar de uma sessão de bate-papo educacional. Relevância: A EAD contém 15,8% das matrículas de graduação no Brasil e o bate-papo é o terceiro meio de interação <i>online</i> mais utilizado, atrás apenas do e-mail e do fórum.
Artefato (Seção 4)	O artefato produzido pela pesquisa é um modelo matemático para estimar o número de mensagens enviadas em função da quantidade de participantes na sessão de bate-papo. Esse modelo é a base para responder a questão da pesquisa, que é determinar a quantidade de pessoas que devem participar de uma sessão de bate-papo educacional.
Processo de Busca da Solução (Seção 5)	Método de Design: A construção do artefato foi realizada utilizando a Modelagem Matemática com Teoria de Filas. O modelo de participação em bate-papo foi desenvolvido num processo iterativo em que, a cada iteração, o modelo foi alterado com base nas avaliações realizadas.
Rigor da Pesquisa (Seções 6)	Método de Avaliação: Pesquisa de caráter quantitativo em que foram comparadas as medidas obtidas com o modelo e as medidas extraídas de sessões reais de bate-papo realizadas num curso a distância, visando avaliar adequação do modelo.
Avaliação (Seção 6)	A avaliação do modelo foi baseada na comparação dos resultados obtidos do modelo e os dados extraídos de 142 sessões reais de bate-papo realizadas num curso a distância. Considerando o Teorema Central do Limite, 95% dos valores reais deveriam estar contidos a uma distância de duas vezes o desvio padrão (para mais e para menos) do valor estimado pelo modelo para o número de mensagens enviadas pelos participantes – conforme discutido na Seção 6. E de fato esse valor foi obtido, o que é um indício da adequação do modelo.
Contribuições da Pesquisa (Seção 7)	As contribuições da pesquisa são: um modelo da quantidade de mensagens enviadas em função da quantidade de participantes de uma sessão de bate-papo educacional; e um maior entendimento sobre a participação em bate-papo educacional.
Comunicação da Pesquisa (Seção 7)	A comunicação da pesquisa visa o público técnico de Modelagem de Sistemas de Computação e Comunicação (por descrever o processo de condução da pesquisa) e também o público responsável pela gestão de cursos de graduação a distância (por disponibilizar um artefato para dimensionar a quantidade de alunos em sessões de bate-papos educacionais).

A abordagem Design Science Research orienta que a pesquisa deve ser conduzida pela aplicação de métodos rigorosos tanto na construção quanto na avaliação do artefato. Na presente pesquisa, buscou-se seguir as diretrizes propostas pela Design Science Research e o desenvolvimento é apresentado nas seções seguintes.

3. Relevância do Problema

O objetivo das pesquisas em Sistemas de Informação é produzir conhecimento que possibilite o desenvolvimento e implementação de soluções baseadas em tecnologias para problemas importantes de negócio até agora não resolvidos [Hevner et al. 2004]. Como

apresentado na Introdução, o número de alunos que optam por realizar sua graduação na modalidade a distância tem aumentado a cada ano e meios de conversação, como o fórum e o bate-papo, são muito usados na educação *online*. Hratinski (2008a) defende que o uso de sistemas síncronos em conjunto com sistemas assíncronos contribui para o aumento da participação dos alunos na educação *online*. Silva (2013) afirma que interface de bate-papo é o espaço virtual mais próximo da sala de aula tradicional e, por ser síncrono, permite a construção do sentimento de pertencimento. Apesar da grande popularidade e reconhecida utilidade, é preciso gerar mais conhecimento sobre o uso do bate-papo no contexto educacional. Na presente pesquisa, buscou-se equacionar quantos alunos podem participar de uma sessão de bate-papo sem decair o nível de participação esperado pelo professor.

Stahl (2009) chama atenção para o impacto do tamanho do grupo sobre a eficácia dos diferentes meios de comunicação; ele defende que grupos pequenos, de três ou quatro alunos ativos, são mais complexos e interessantes do que uma dupla respondendo um ao outro. Contudo, se o grupo for grande demais ocorrerão problemas. A “confusão no bate-papo” [Fuks et al. 2006] ocorre quando os participantes não conseguem acompanhar a discussão em função de fenômenos como, por exemplo, a Sobrecarga de Mensagens, causada pela grande quantidade de mensagens postadas em um curto intervalo de tempo decorrente de um número excessivo de participantes na sessão de bate-papo.

Alguns autores, a partir de dados empíricos e em contextos específicos, indicam que o número ideal de participantes num bate-papo é entre 4 e 5 pessoas [Böhlke 2003]; há quem defenda que o bate-papo é mais produtivo com 7 ou mais participantes [Löber et al. 2007]; há quem defenda que o grupo deve ser inferior a 10 participantes [Hrastinski 2008b]; há quem indique que 10 é um tamanho bom para o grupo de participantes [Piva et al. 2011]; e há quem indique que a quantidade de participantes deve estar entre 6 a 18 alunos [Pimentel 2003]. Alguns sistemas de bate-papo limitam o número de participantes a 30 usuários por sessão, como é o caso do Bate-papo UOL; e há outros sistemas que não estabelecem limites. Qual desses números é o mais correto?

4. Artefato: Modelo de Participação em Bate-papo Educacional

Nas pesquisas em Sistemas de Informação, desenvolve-se um artefato visando resolver um problema relevante. Na presente pesquisa, foi desenvolvido um modelo matemático para estimar quantos alunos devem participar de uma sessão de bate-papo educacional.

Num bate-papo educacional, diferentemente do recreativo, é importante que todos consigam ler todas as mensagens e que também possam enviar mensagens de vez em quando. Se não fosse essa restrição, não haveria um fator limitante para a quantidade de pessoas no bate-papo. Só faz sentido discutir a quantidade máxima de participantes se for exigido um nível mínimo de participação, pois caso a participação possa ser zero (o aluno não precisa ler nem escrever mensagens), tal como ocorre na transmissão *broadcasting*, então não há um limite máximo de alunos que devem participar da sessão, e o professor pode dar uma aula por bate-papo para milhares de alunos ao mesmo tempo desde que somente o professor esteja digitando mensagens e a participação dos alunos tenda a zero. O problema surge quando se dá a chance para o aluno interagir com o professor e com os colegas – e é para isso que serve o bate-papo. No contexto educacional, não é razoável colocar centenas de alunos conversando ao mesmo tempo, por isso deve-se determinar a quantidade máxima de participantes numa sessão de bate-

papo para que seja garantido que o aluno consiga ler todas as mensagens e que seja possível escrever uma determinada quantidade de mensagens.

O número total de mensagens produzidas na sessão aumenta proporcionalmente à quantidade de participantes; conforme aumenta a quantidade de mensagens, o participante gasta mais tempo para ler todas as mensagens, e terá menos tempo para enviar novas mensagens, e assim sua produção de mensagens decai com o aumento do número de participantes. O objetivo do modelo desenvolvido nessa pesquisa é equacionar a produção de mensagens em função do tamanho do grupo, para que seja possível estimar o limite máximo de participantes em que é possível garantir o mínimo de participação exigido pelo professor.

O participante escreve uma mensagem e envia para os outros participantes que, por sua vez, irão ler as mensagens recebidas e interagir escrevendo novas mensagens. Quando o participante está lendo ou escrevendo alguma mensagem e nesse momento chegam outras mensagens a serem lidas, essas são enfileiradas para a posterior leitura. Desta forma, para elaborar o modelo foi empregada a Teoria de Filas [Kleinrock 1975] – Figura 3 – em que cada participante é representado como um servidor que precisa ler (servir) todas as mensagens (clientes) digitadas pelos outros participantes, que vão sendo enfileiradas para a leitura na ordem em que são recebidas. Cada participante, além de ser um servidor, também desempenha o papel de uma fonte geradora de mensagens que são enviadas para as filas dos demais participantes.

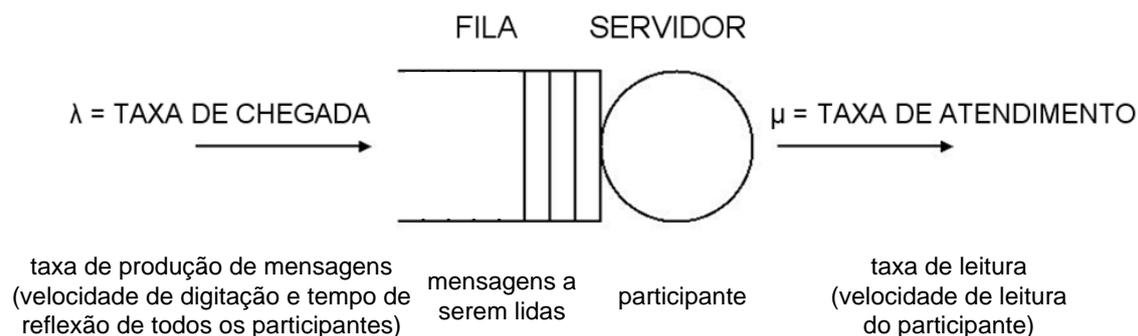


Figura 3. Sistema com fila única

Dentre os parâmetros desse modelo, identifica-se a taxa de leitura de mensagens μ , que é o inverso do tempo médio de leitura das mensagens produzidas na sessão; e a taxa de produção de mensagens λ , que se deseja determinar dado o tamanho p do grupo de participantes. Os tempos de leitura e de produção de mensagens são considerados, na presente pesquisa, como exponencialmente distribuídos.

Para obter as estimativas de produção de mensagem a partir do modelo construído, o modelo foi implementado no Tangram-II [De Souza e Silva et al. 2009]. O Tangram-II é um ambiente de modelagem baseado no paradigma orientado a objetos que fornece uma variedade de *solvers* para obtenção das medidas de interesse. As estimativas produzidas a partir do modelo elaborado na presente pesquisa foram obtidas a partir da Solução Analítica do Tangram-II.

5. Processo de Busca de Solução

O ciclo de design (construção/avaliação) é o coração de todo projeto de pesquisa em Ciência de Design e também onde é feito o trabalho mais intenso em Design Science

Research [Hevner 2007]. Deve-se descrever todo o processo de condução da pesquisa e deve-se aplicar métodos rigorosos tanto na construção quanto na avaliação do artefato. O rigor da pesquisa é derivado do uso eficaz da base de conhecimento, que é o conjunto de fundamentos e metodologias necessárias para a realização da pesquisa.

A construção do artefato é feita por meio de um processo iterativo. A cada iteração, os dados produzidos pelo modelo foram comparados com os dados extraídos de cenários reais de bate-papo na educação, duas aulas sobre linguagem de programação Java e uma aula sobre acessibilidade – 14, 6 e 9 participantes respectivamente, todas com alunos de graduação em Informática. Em seguida o artefato foi refinado visando obter estimativas mais precisas. Esse é o típico procedimento adotado no *design* de soluções, em que a atividade de *design* oscila entre a atividade conceitual e prática, como esquematizado na Figura 4.

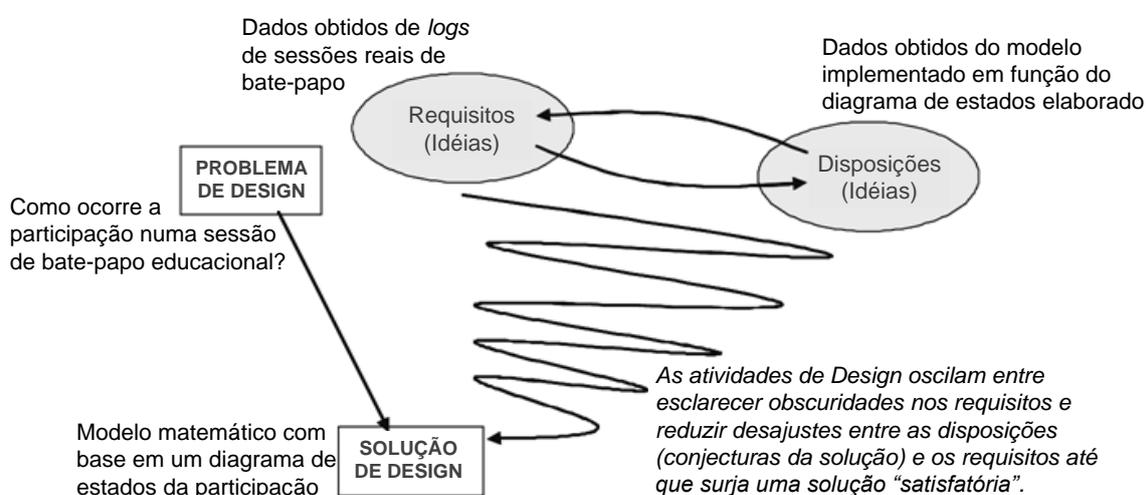


Figura 4. Design como processo iterativo em busca de uma solução para um problema [Mckay et al., 2008, p.5]

Para avaliar se o modelo produz estimativas com uma precisão dentro de um limite de aceitação, essas estimativas foram comparadas com dados extraídos de *logs* reais de bate-papo das turmas de graduação. Dos *logs* de bate-papo, foram extraídas a duração do bate-papo, a quantidade de participantes e a quantidade de mensagens enviadas. Espera-se que o modelo, com os parâmetros da turma, produza estimativas próximas dos dados reais. Isto é avaliado na próxima seção.

5.1. Ações realizadas durante uma sessão de bate-papo

Antes de partir para a modelagem da participação, buscou-se entender melhor as ações realizadas pelos participantes durante uma sessão de bate-papo. Observou-se que o participante lê e escreve mensagens, e fica refletindo em alguns momentos. Essas ações, descritas nas subseções a seguir, são os comportamentos elementares de qualquer participante de um bate-papo educacional e constituem a base do modelo matemático apresentado na presente pesquisa.

Leitura

Pesquisas apontam que o processo de leitura sofre influência dos componentes estruturais do texto. Algumas unidades estruturais como letras e palavras familiares, são

incorporadas (*overlearned*) após experiência de leitura prolongada. Estudos sugerem que esses componentes microestruturais diferenciam leitores rápidos e leitores lentos [Graesser et al., 1980]. Estudos mais recentes, como o Modelo de Componentes da Leitura [Aaron et al., 2008], explicam de maneira mais ampla a competência em leitura como sendo resultado da atuação e integração de três domínios: cognitivo (reconhecimento de palavras e compreensão), psicológico (aspectos motivacionais e emocionais relacionados à aprendizagem) e ecológico (aspectos familiares, culturais e escolares).

Em uma sessão de bate-papo, uma pessoa não lê igual a outra pessoa; a velocidade de leitura depende do nível de escolaridade, da experiência em bater-papo, do grau de interesse e de conhecimento sobre o assunto em discussão, dentre outras características do indivíduo. O tempo de leitura também depende de fatores externos ao sujeito, como o tamanho da mensagem e a interface do sistema de bate-papo. Esses fatores também estão relacionados aos domínios (cognitivo, psicológico e ecológico).

Para caracterizar o tempo de leitura de mensagens de bate-papo, foi realizado um experimento com 6 alunos de pós-graduação em Informática, sendo 4 homens e 2 mulheres com idade entre 24 e 55 anos. Para medir o tempo de leitura, cada indivíduo leu mensagem-a-mensagem de um *log* real de bate-papo educacional pela tela do computador. Quando o indivíduo pressionava uma tecla, uma nova mensagem era apresentada e era registrado o tempo de leitura decorrido entre o pressionar da tecla. Buscou-se caracterizar como ocorre a leitura de mensagens entre diferentes indivíduos ao ler mensagens de diferentes tamanhos, em *logs* sobre diferentes assuntos e apresentados em diferentes interfaces de bate-papo. O estudo encontra-se detalhado em [Rocha 2013]. Um dos principais achados é que o tempo de leitura é proporcional à quantidade de caracteres da mensagem (correlação de 0,72), o que possibilita interpretar que as variáveis tem forte correlação [Shimakura, 2006]. A partir desse resultado foi assumido, por simplificação, que o tempo de leitura (\bar{T}_L), medido em segundos, pode ser modelado por uma equação linear em função da quantidade de caracteres da mensagem conforme apresentado na equação (1).

$$\bar{T}_L = 1 + (0,04 * \text{média de caracteres por mensagem}) \quad (1)$$

A equação (1) representa o tempo médio, em segundos, que os sujeitos do experimento dispenderam para a leitura das mensagens daquelas sessões de bate-papo (calculada a partir das linhas de tendência linear da correlação entre o tempo de leitura e a quantidade de caracteres por mensagem por indivíduo).

Escrita

Cada pessoa possui uma velocidade de digitação específica. Não só a velocidade de digitação, mas também os padrões de digitação diferem de um indivíduo para outro. O intervalo entre o pressionar de uma tecla e a outra, e o tempo em que uma tecla permanece pressionada, são características únicas de cada indivíduo e corroboram para o estabelecimento de um padrão de digitação individual [Gaines et al., 1980]. Esses padrões são estudados na área de Biometria da Digitação.

Para a presente pesquisa, é de interesse obter tempos médios de digitação de mensagens durante a participação em bate-papos educacionais, sendo preciso medir a velocidade de elaboração de mensagem dos participantes. Para efetivar essa medição, foi desenvolvido um mecanismo que captura o tempo de cada tecla pressionada durante a

digitação de uma mensagem de bate-papo. Foi realizado um estudo, com profissionais de computação, em que os sujeitos digitaram algumas mensagens e o tempo de digitação foi registrado e posteriormente analisado. Nesse experimento, o tamanho médio das mensagens digitadas foi de 85 caracteres por mensagem e o tempo médio de digitação foi de 24 segundos por mensagem, resultando na velocidade média de digitação de 1 caractere / 0,28 segundos. Verificou-se uma correlação muito forte (0,95) entre a quantidade de caracteres e o tempo de digitação da mensagem. A partir desse resultado foi assumido que o tempo de escrita (\bar{T}_E) pode ser modelado por uma equação linear em função da quantidade de caracteres da mensagem.

$$\bar{T}_E = 0,28 * \text{média de caracteres por mensagem} \quad (2)$$

A equação (2) descreve o tempo médio de escrita de mensagens, medido em segundos, dispendido pelos sujeitos do grupo do experimento.

Reflexão

Os participantes não ficam lendo e digitando o tempo todo durante a sessão de bate-papo. Ocorrem intervalos, por exemplo, entre o término da leitura das mensagens publicadas até aquele momento e a chegada de uma nova mensagem a ser lida, em que o participante não está digitando nem lendo mensagens. Nestes momentos o usuário encontra-se refletindo sobre a discussão. O tempo de reflexão é o tempo que sobra de uma sessão de bate-papo desconsiderando o tempo em que o usuário fica escrevendo suas mensagens e o tempo em que fica lendo as mensagens. Assim, nesta pesquisa, o tempo de reflexão foi estimado por meio da equação (3).

$$T_R = \text{duração da sessão} - (\bar{m} * \bar{T}_E) - (\bar{m} * p * \bar{T}_L) \quad (3)$$

onde T_R é o tempo de reflexão numa sessão de bate-papo, \bar{m} é a quantidade média de mensagens enviadas por participante na sessão, e p é a quantidade de participantes.

5.2. Modelagem do comportamento de um usuário de bate-papo

Para modelar o comportamento de um participante da sessão de bate-papo educacional, foi desenvolvido um modelo de estados para representar as atividades primárias que o participante realiza durante o bate-papo.

Escrita-Leitura

O modelo de estados mais simples é supor que o participante fica lendo ou escrevendo mensagens durante a sessão de bate-papo (nunca fica ocioso). Para esse caso, o modelo de estados possui apenas dois estados conforme proposto em [Rocha et al., 2012] e ilustrado na Figura 5.

Na primeira tentativa de modelar o comportamento do participante, este permanece no estado Escrevendo por um determinado tempo médio até que envia a mensagem aos demais participantes. Após enviar a mensagem, ele verifica se existem mensagens a serem lidas, e se houver, passa para o estado Lendo. Enquanto o participante está Escrevendo ou Lendo mensagens, outras mensagens podem chegar e essas são enfileiradas na ordem em que são enviadas pelos demais participantes em um Fila de Mensagens. A transição de estados é modelada em função de taxas. O participante permanece Lendo uma determinada mensagem por um tempo médio de leitura (\bar{T}_L), e a taxaLeitura é calculada como o inverso do tempo médio. Da mesma forma, o

participante permanece no estado Escrevendo durante um tempo médio de escrita (\bar{T}_E), a partir do qual se calcula a taxa de escrita.

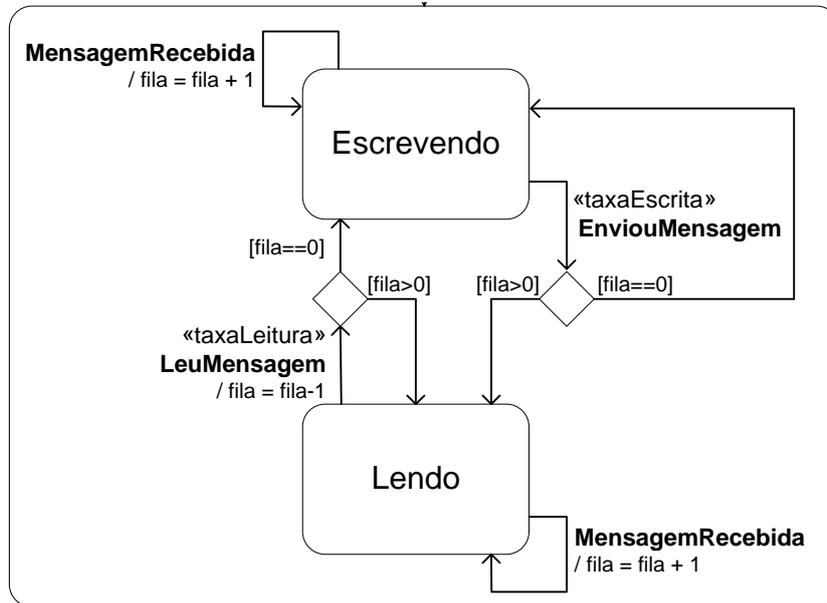


Figura 5. Modelo dos Estados Lendo e Escrevendo

Representar um participante de bate-papo por apenas dois estados é irreal, pois pressupõe que o participante fica o tempo todo ou Lendo ou Escrevendo mensagens sem nunca ficar ocioso. Contudo, essa simplificação da realidade é o que possibilita estimar um Limite Máximo de mensagens produzidas pelo participante numa sessão de bate-papo, pois pressupõe que esse sujeito se encontra o tempo todo ocupado, produzindo o máximo de mensagens que seria capaz. Esse limite é calculado pela equação (4):

$$m = 60 / (\bar{T}_E + \bar{T}_L (p-1)) \quad (4)$$

onde m representa o número de mensagens produzidas por participante em uma sessão de 60 minutos, \bar{T}_E o tempo médio de escrita dos participantes, \bar{T}_L o tempo médio de leitura e p o número de participantes na sessão de bate-papo.

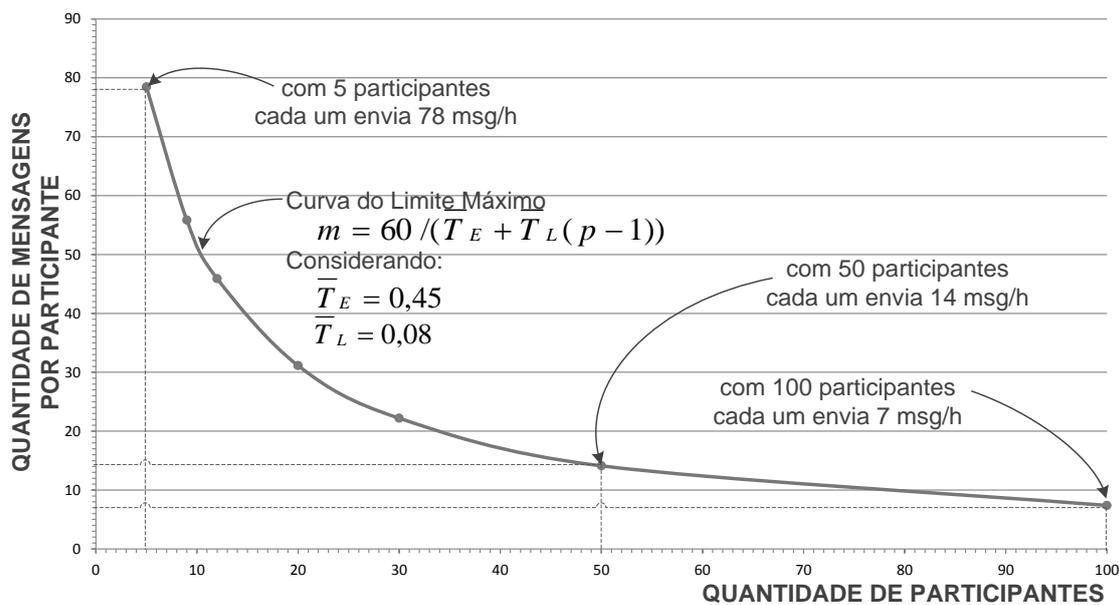


Figura 6. Curva do Limite Máximo

Na Figura 6, é ilustrada a curva de participação obtida pela equação (4) para $\bar{T}_E = 25$ segundos (0,45 min) e $\bar{T}_L = 4,8$ segundos (0,08 min). Para $p = 5$ tem-se uma produção de 78 mensagens, enquanto para $p = 100$, tem-se uma produção de 7 mensagens por participante.

Ao se comparar os dados estimados pelo modelo com dados reais de bate-papo observou-se que a taxa de produção de mensagens estimada pelo modelo era muito superior ao real. Com essa primeira modelagem foi possível constatar que o usuário não pode ficar o tempo todo lendo e escrevendo na sessão, ou seja, essa simplificação não é possível. Em alguns momentos encontra-se ocioso e assumimos que nesse período o participante fica refletindo sobre a sessão, conforme apresentado na Seção 5.1. O entendimento de haver um tempo de reflexão (sem ler nem escrever) deu origem à modificação do modelo, conforme discutido a seguir.

Escrita-Leitura-Reflexão

Uma nova abordagem foi desenvolvida para reduzir a diferença entre o real e o estimado pelo modelo. No segundo modelo de estados, apresentado na Figura 7, leva-se em conta que um participante, além de ler e escrever mensagens, também fica refletindo em alguns momentos.

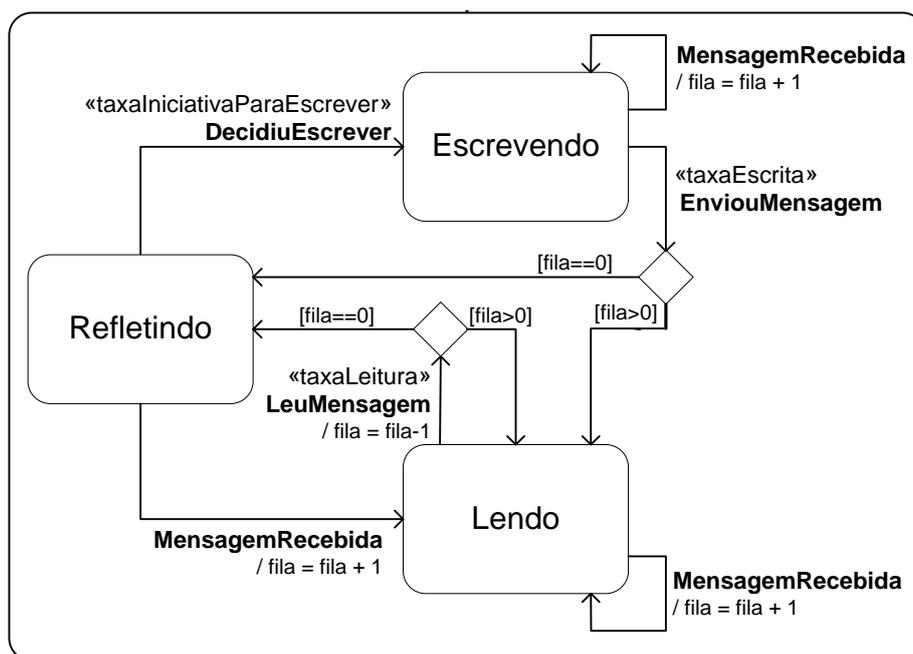


Figura 7. Modelo dos Estados Lendo, Escrevendo e Refletindo

Nessa abordagem, o indivíduo vai para o estado Refletindo quando não tem mais mensagens a serem lidas, e fica refletindo por um tempo até que: chega uma nova mensagem a ser lida e, imediatamente, o usuário volta para o estado Lendo; ou fica refletindo por algum tempo até que decide escrever uma nova mensagem, indo para o estado Escrevendo.

O modelo com estes três estados produz estimativas compatíveis com os dados reais utilizados para validação do modelo (apresentado na Seção 6). Contudo, nesta pesquisa chegou-se a desenvolver um modelo ainda mais refinado contendo quatro estados e explicado a seguir.

Escrita-Leitura-Reflexão-Motivação

Com a intenção de refinar ainda mais o modelo de estados da participação numa sessão de bate-papo, foi considerado o princípio geral “fala um de cada vez” do modelo da organização da conversação [Sacks et al., 1974]. O participante não escreve mensagens aleatoriamente, mas sim busca responder a alguma mensagem enviada anteriormente dando continuidade à conversação. Esse comportamento foi representado pela motivação do participante em escrever novas mensagens: o participante precisa estar motivado para escrever uma nova mensagem, e essa motivação só ocorre quando o participante lê alguma mensagem de outro participante. Por outro lado, a motivação não se aplica ao estado Lendo, pois conforme apresentado na Seção 4, entende-se que a participação é medida pela leitura de todas as mensagens postadas pelos demais participantes e pelo envio de mensagens em resposta às mensagens lidas pelo participante. Com isso, o participante nunca pode estar “desmotivado para ler”.

Considerando a motivação do participante para escrever mensagens e a ausência dessa motivação, foi elaborado o modelo de estados apresentado na Figura 8.

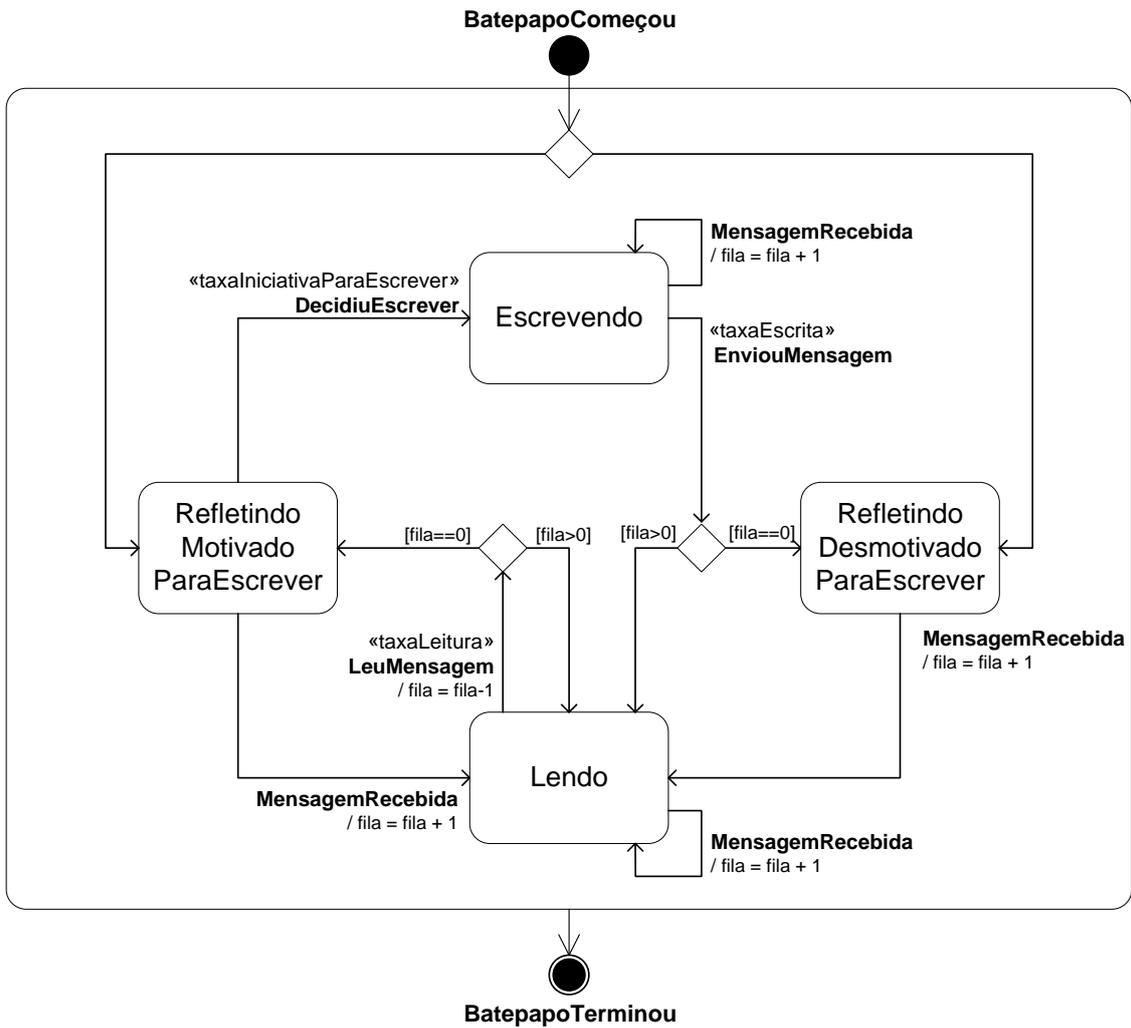


Figura 8. Modelo dos Estados desmembrando Refletindo (Motivado e Desmotivado)

O modelo baseado em quatro estados somente produz estimativas diferentes do modelo baseado em três estados quando há poucos participantes (até 5 ou 6 participantes), sendo observado um aumento discreto e gradativo na produção de mensagens até essa quantidade de participantes. Ao se incrementar mais participantes ao grupo de bate-papo, a produção de mensagens por participante começa a decair conforme ilustrado na Figura 9.

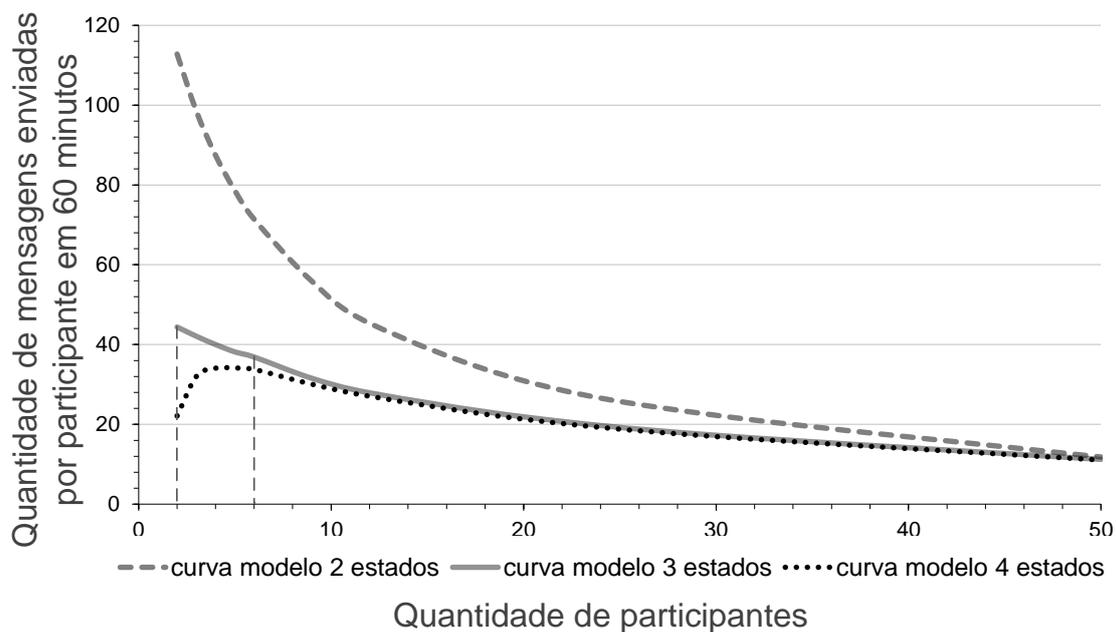


Figura 9. Comparação entre os modelos com 2, 3 e 4 estados para um mesmo contexto

As sessões reais de bate-papo usadas como referência para a comparação não contêm tão poucos participantes, e assim não foi possível investigar qual dos dois modelos produz estimativas mais precisas. Sendo assim, voltamos um passo e adotamos o modelo apresentado na Figura 6 por ser mais simples. Esse processo iterativo em busca da construção do artefato que resolva o problema investigado é característico de Design Science Research conforme já discutido no início desta seção.

5.3. Modelagem da interação entre os participantes da sessão de bate-papo

Conforme apresentado na Seção 4, para se obter as estimativas de produção de mensagens, o modelo foi implementado no ambiente de modelagem Tangram-II, que é baseada no paradigma orientado a objetos. A seguir são apresentadas as duas estratégias de modelagem desenvolvidas na pesquisa e suas respectivas vantagens e desvantagens.

Estratégia de Modelagem: Interação entre Indivíduos

Na primeira estratégia de modelagem da interação entre os participantes de um bate-papo educacional, o modelo é formado por um grupo de objetos do tipo “**Participante**” que possui o comportamento descrito no Modelo de Estados da Figura 6. O objeto Participante é então replicado p -vezes até se representar todos os participantes do grupo. A essa estratégia de modelagem foi dado o nome Interação entre Indivíduos e conceitualmente é representada na Figura 10.

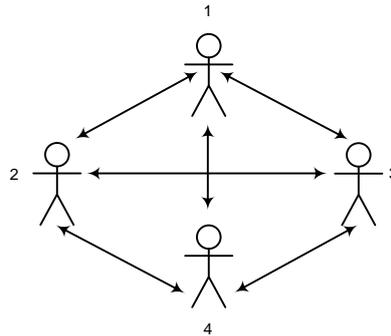


Figura 10. Estratégia de Modelagem: Interação entre Indivíduos

Após o modelo estar implementado no Tangram II (módulo *Model Specification*), gera-se a Cadeia de Markov contendo o conjunto de todos os estados dos p participantes bem como as transições possíveis entre os estados (módulo *Mathematical Model*). Segundo Fogliatti e Mattos (2007) “Cadeia de Markov é um processo estocástico com espaço de estados discretos, onde a cadeia é formada por um conjunto de variáveis e a probabilidade do próximo estado depende apenas do estado atual – memoryless”. Por fim, as estimativas de produção de mensagens por participante são obtidas pela técnica de análise transiente *Expected Cumulative Impulse Reward* (módulo *Analytical Model Solution*). A cada envio de mensagem pelo participante, uma recompensa (*reward*) é adicionada. Recompensas são recursos usados para calcular uma série de medidas de interesse tanto analiticamente quanto por meio de simulação [De Souza e Silva et al. 2009, p. 65].

Por meio da estratégia de modelagem Interação entre Indivíduos, só é possível obter as estimativas de mensagens para até 5 participantes, pois com 6 ou mais participantes ocorre uma explosão de estados já que o número de estados da Cadeia de Markov aumenta exponencialmente a medida que o número de participantes é incrementado. Como um dos objetivos da pesquisa é estimar o máximo de participantes que devem participar de uma sessão de bate-papo mantendo uma determinada produção de mensagens, essa solução se mostrou inviável.

Estratégia de Modelagem: Interação Indivíduo-Grupo

Continuando o processo de Busca da Solução, para contornar o problema de explosão de estados do ambiente Tangram-II, foi elaborada uma segunda estratégia de modelagem com apenas 2 objetos: um **Participante**, com o comportamento descrito no modelo de estados da Figura 7, e um **Grupo de Participantes** modelado como uma fonte geradora de mensagens, que produz mensagens para o Participante com uma taxaEnvioMsg – Figura 11.

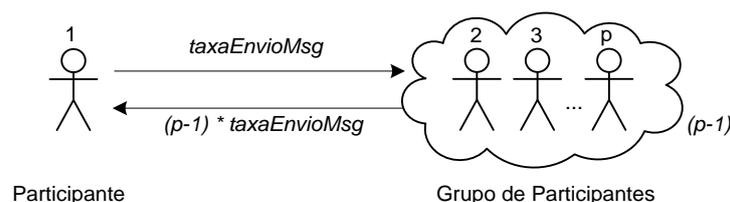


Figura 11. Estratégia de Modelagem: Interação Indivíduo-Grupo

Nessa estratégia, cada participante do Grupo produz mensagens numa taxa semelhante ao do indivíduo isolado, pois assume-se que todos os participantes do bate-papo têm comportamento similar. Isto significa que a taxa de envio do Grupo dividida por $(p-1)$ precisa ser próxima ao número de mensagens enviadas pelo indivíduo isolado. O valor da taxa de mensagens enviadas pelo indivíduo isolado é a medida de interesse que se quer encontrar. Para obter essa taxa, o primeiro passo é atribuir um valor aleatório para a taxa de produção de mensagens pelo grupo e, pelo Tangram-II, calcular a taxa de envio de mensagens do indivíduo isolado. Em seguida compara-se a taxa atribuída ao grupo e a produzida pelo indivíduo. A diferença entre as taxas deve ser a menor possível - na presente pesquisa, foi definida que a diferença deveria ser inferior a 1 mensagem. Iterativamente vai se ajustando a estimativa da taxa de envio de mensagens do grupo até que a taxa dos indivíduos do grupo fique próxima da taxa do indivíduo isolado. Desta forma foi possível obter, com o Tangram II, as estimativas para a medida de interesse da presente pesquisa para qualquer tamanho de grupo. Como exemplo, na Tabela 2 são exemplificados os passos realizados iterativamente para se obter uma estimativa em que o envio de mensagens do grupo e do participante seja menor do que 1 mensagem.

Tabela 2. Obtenção da estimativa de mensagens enviadas adotando a estratégia de Interação Grupo-Indivíduo

Sessão de bate-papo de 60 minutos com 7 participantes ($p = 7$)	
“Chute” inicial: Taxa de envio de mensagens do Grupo e a respectiva quantidade de mensagens enviadas por cada participante do Grupo (mensagens por indivíduo do grupo durante 1h de sessão = taxa * 60min /6 participantes)	Valor calculado pelo Tangram-II: quantidade de mensagens enviadas pelo Participante Isolado durante 1h de bate-papo
3,3 (33,0)	34,2
3,4 (34,2)	34,0
3,6 (36,0)	33,6

Conforme apresentado na Tabela 2, foram testadas as sucessivas taxas de envio de mensagens do Grupo: 3,3; 3,4 e 3,6 mensagens por minuto. Dentre as taxas testadas, a que mais aproximou a quantidade de mensagens enviadas por cada participante do grupo com a do participante isolado foi 3,4. Para essa taxa, cada participante do grupo produziu 34,2 mensagens, enquanto o participante isolado enviou 34,0 mensagens durante 60 minutos. Essa diferença é inferior a 1 mensagem, e restringindo a precisão da taxa a uma casa decimal, essa é a taxa que mais aproxima a produção do indivíduo isolado com a do indivíduo do grupo. Repetindo esse trabalho para diferentes quantidades de participantes no grupo, obtém-se a medida de interesse: quantidade de mensagens produzidas pelo participante da sessão de bate-papo em função do tamanho do grupo.

A combinação do modelo de 3 estados (Figura 7) e a estratégia de modelagem Interação Indivíduo-Grupo (Figura 11) possibilitou a obtenção de estimativas próximas às encontradas em bate-papos reais. De posse da solução (artefato) para o problema de estimativa da quantidade de participantes de um bate-papo educacional, encerra-se o ciclo construção-avaliação da Design Science Research para essa pesquisa. Em seguida, é feita a avaliação final do modelo.

6. Rigor da Pesquisa: Avaliação do Modelo

Diferente do processo de busca da solução, em que o artefato (modelo) era construído e em seguida avaliado comparando-se as estimativas produzidas com os dados obtidos dos 3 *logs* reais (Seção 5), a avaliação final do artefato foi feita pela comparação das estimativas produzidas pelo modelo com os dados reais obtidos de *logs* de bate-papo das turmas de uma disciplina lecionada totalmente a distância, denominada “Tecnologias de Informação Aplicadas à Educação” (TIAE) [Lucena e Fuks 2002].

De acordo com Wieringa (2014), respostas a questões de conhecimento, ou fenômeno comportamental segundo Hervner et al. (2004), podem ser obtidas tanto por meio de pesquisa empírica, com a aplicação do artefato em um contexto, quanto por meio de pesquisa analítica, com a aplicação de métodos matemáticos. Questões de conhecimento analíticas podem ser respondidas por meio de análise conceitual, como matemática ou lógica, sem a coleta de dados do mundo real [Wieringa, 2014, pg. 17]. Na presente pesquisa, foi feita a avaliação analítica do artefato com a aplicação do Teorema Central do Limite.

Ao todo foram obtidos 142 *logs* de bate-papo ocorridos em 18 turmas ministradas entre os anos de 1999 e 2008 (em cada turma foram realizadas, em média, 8 sessões de bate-papo). Para realizar a avaliação do modelo, foram extraídas dos *logs* de bate-papo das turmas TIAE as seguintes métricas de interesse: duração da sessão de bate-papo, quantidade de participantes, quantidade de mensagens enviadas por participante e quantidade média de caracteres por mensagem.

A parametrização do modelo de participação foi feita com base nos estudos apresentados na subseção (5.1) sobre o processo de leitura e o processo de escrita de mensagens em bate-papo educacional. Como apresentado na subseção 5.1, foi identificada uma alta correlação entre o tamanho médio da mensagem e os tempos médios de leitura e tempo médio de escrita. Com isso, a quantidade média de caracteres por mensagem extraída dos *logs* reais serviu de base para calcular os tempos de leitura e escrita. Com a informação da duração média de cada turma e da produção média de mensagens por aluno, foi possível inferir o tempo médio de reflexão dos alunos. Como o modelo todo parametrizado, as estimativas de produção de mensagens para cada tamanho de grupo puderam ser obtidas pelo Tangram-II.

Os parâmetros do modelo são: *taxaLeitura*, *taxaEscrita* e *taxaIniciativaParaEscrever* – esses parâmetros indicam a frequência de ocorrência de cada um dos eventos que compõem os estados do modelo ilustrado na Figura 4. Para parametrizar o modelo, é necessário medir o tempo de leitura, tempo de digitação e tempo de iniciativa para escrita no contexto a ser aplicado. Essas medições podem ser realizadas em uma pré-sessão de bate-papo com o grupo em questão. Além dessas taxas, também é preciso especificar a função de distribuição de probabilidade de cada evento.

Uma vez que não é possível voltar no tempo para medir a velocidade de leitura e de digitação dos participantes das sessões de bate-papo que compõem o corpus de análise, então optou-se por estimar os parâmetros de cada turma TIAE com base nas equações apresentadas na Seção 5. Considerou-se que o tempo médio de leitura da equação (1) é uma aproximação razoável para o contexto das turmas TIAE, pois ambos os contextos são compostos de alunos de graduação e pós-graduação em Informática. Para estimar o parâmetro *taxaLeitura* das turmas TIAE, além da velocidade de digitação também é preciso considerar o tamanho médio das mensagens emitidas em cada turma,

sendo este um parâmetro obtido dos *logs*. Na equação (5), é exemplificada a aplicação da equação (1) para estimar o parâmetro taxaLeitura dos participantes da turma TIAE2007.1 (primeiro semestre de 2007), onde:

$$\begin{aligned}\bar{T}_L &= 1 + (0,04 * \text{média de caracteres}) \\ &= 1 + (0,04 * 94,1) = 1 + 3,76 = 4,76 \text{ seg} = 0,08 \text{ min} \\ \text{taxaLeitura} &= 1 / \bar{T}_L = 1/0,08 = 12,6 \text{ msg/min}\end{aligned}\quad (5)$$

Para estimar as taxas de escrita dos participantes das turmas TIAE, na impossibilidade de se obter esse parâmetro a partir do *log*, considerou-se que a equação (2) é uma aproximação razoável. Esta equação foi parametrizada com o tamanho médio das mensagens de cada turma TIAE (obtido dos *logs* reais). Na equação (6), é exemplificada a aplicação da equação (2) para estimar o parâmetro taxaEscrita dos participantes da turma TIAE2007.1, onde:

$$\begin{aligned}\bar{T}_E &= 0,28 * \text{média de caracteres por mensagem} \\ &= 0,28 * 94,1 = 26,3 \text{ seg} = 0,44 \text{ min} \\ \text{taxaEscrita} &= 1/\bar{T}_E = 1/0,44 = 2,3 \text{ msg/min}\end{aligned}\quad (6)$$

Para estimar as taxaIniciativaParaEscrever dos participantes das turmas TIAE, foi primeiramente utilizada a equação (3) para estimar o tempo de reflexão total na sessão de bate-papo (T_R) considerando: duração da sessão (obtida de cada *log*), a quantidade de participantes de cada turma TIAE; a média de mensagens enviadas por participante; e os tempos de leitura e de escrita estimados de cada turma. Considerando-se que o participante envia em média (\bar{m}) mensagens durante a sessão de bate-papo, então a escrita de uma mensagem e a outra esse participante fica refletindo, em média, T_R / \bar{m} . Optou-se em adotar esse valor como uma aproximação do tempo médio de iniciativa de escrita de mensagens (\bar{T}_{IE}). A taxa de iniciativa para escrever é então obtida pelo inverso de \bar{T}_{IE} . O procedimento elaborado para a estimativa desses parâmetros na turma TIAE2007.1 é exemplificado na equação (7).

$$\begin{aligned}T_R &= \text{duração da sessão} - (\bar{T}_L * \bar{m} * p) - (\bar{T}_E * \bar{m}) \\ T_R &= 60 - (0,08 * 32,8 * 7,9) - (0,44 * 32,8) = 24,8 \text{ min} \\ \bar{T}_{IE} &= T_R / \bar{m} \\ \bar{T}_{IE} &= 24,8 / 32,8 = 0,76 \\ \text{taxaIniciativaParaEscrever} &= 1 / \bar{T}_{IE} = 1/0,76 = 1,3\end{aligned}\quad (7)$$

Além das taxas, é preciso especificar a função de distribuição de probabilidade dos eventos. Na presente pesquisa, optou-se por utilizar o modelo Poisson com a distribuição exponencial para representar a distribuição de todos os eventos do modelo já que essa distribuição representa um grande número de processos na natureza e é a distribuição mais comumente utilizada em processos modelados como fila. Além disso, de acordo com Fogliatti e Mattos (2007), "*sempre que possível, deve ser escolhida a distribuição exponencial mesmo que não seja o melhor ajuste, pois a sua utilização em muito facilita a obtenção das medidas de desempenho do sistema*" (p.219).

Com as taxas usadas como parâmetros do modelo considerando os dados de cada

turma TIAE, e especificadas as funções de distribuição de probabilidade dos eventos, o modelo pôde ser rodado no ambiente Tangram-II para obter o número de mensagens enviadas em função das seguintes quantidades de participantes: 5, 9, 12, 20, 30, 50 e 100 participantes. Com os valores calculados pelo modelo e interpolando os valores não calculados, é traçada a “curva de participação” estimada para cada turma TIAE. Como exemplo, na Figura 12 é apresentada a curva traçada a partir dos resultados produzidos pelo modelo para a turma TIAE 2007.1.

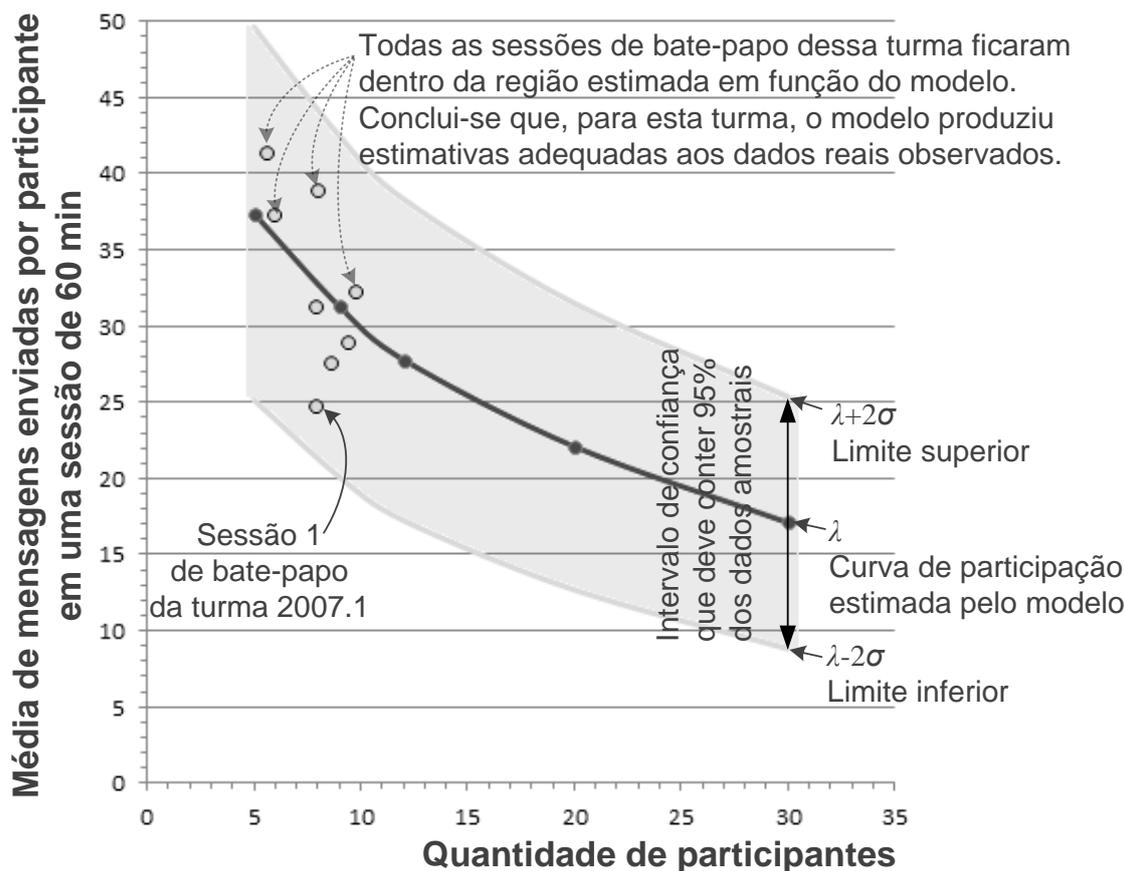


Figura 12. Curva de Participação e Limites Superior e Inferior para Avaliação

A curva de participação é traçada pela união dos pontos estimados pelo modelo de produção de mensagens para cada tamanho do grupo de participantes. Como ilustrado na Figura 8, foi verificado se a quantidade de mensagens produzidas por participante em cada sessão da turma TIAE ficou dentro de certos limites ao redor do ponto estimado pelo modelo. Esses limites são obtidos pelo Teorema Central do Limite, em que conforme se aumenta o tamanho da amostra (ocorrências reais), a distribuição amostral da média aproxima-se de uma distribuição Normal, e neste caso 95% dos dados amostrais ficam entre $\mu-2\sigma$ e $\mu+2\sigma$, onde μ é a média populacional e σ é o desvio padrão populacional. De todas as sessões de bate-papo das turmas TIAE, 95,8% ficaram dentro da região estimada em função do modelo. Este valor é bem próximo ao valor esperado pelo Teorema Central do Limite (95,4%), o que possibilita concluir que o modelo de participação produz estimativas adequadas da média de mensagens enviadas por sessão de bate-papo em função do tamanho do grupo.

Uma vez que o modelo foi considerado válido, foi possível buscar uma resposta

para a pergunta inicial desta pesquisa: quantos alunos devem participar de uma sessão de bate-papo educacional?

A primeira constatação sobre essa questão é que não existe uma resposta única, pois o limite de alunos depende do contexto de cada turma: do perfil dos alunos (velocidade de leitura e de digitação) e das características da sessão de bate-papo (interface do sistema, protocolo de conversação etc.). No modelo, essas características são encapsuladas nos parâmetros tempo médio de escrita, leitura e reflexão. Estimados esses parâmetros para uma turma, ou para o perfil médio de alunos de um curso, o modelo é usado para traçar a curva de participação (Figura 13). Além do contexto específico de cada turma, o limite de alunos também depende da expectativa da quantidade de mensagens que cada aluno deve produzir durante a sessão.

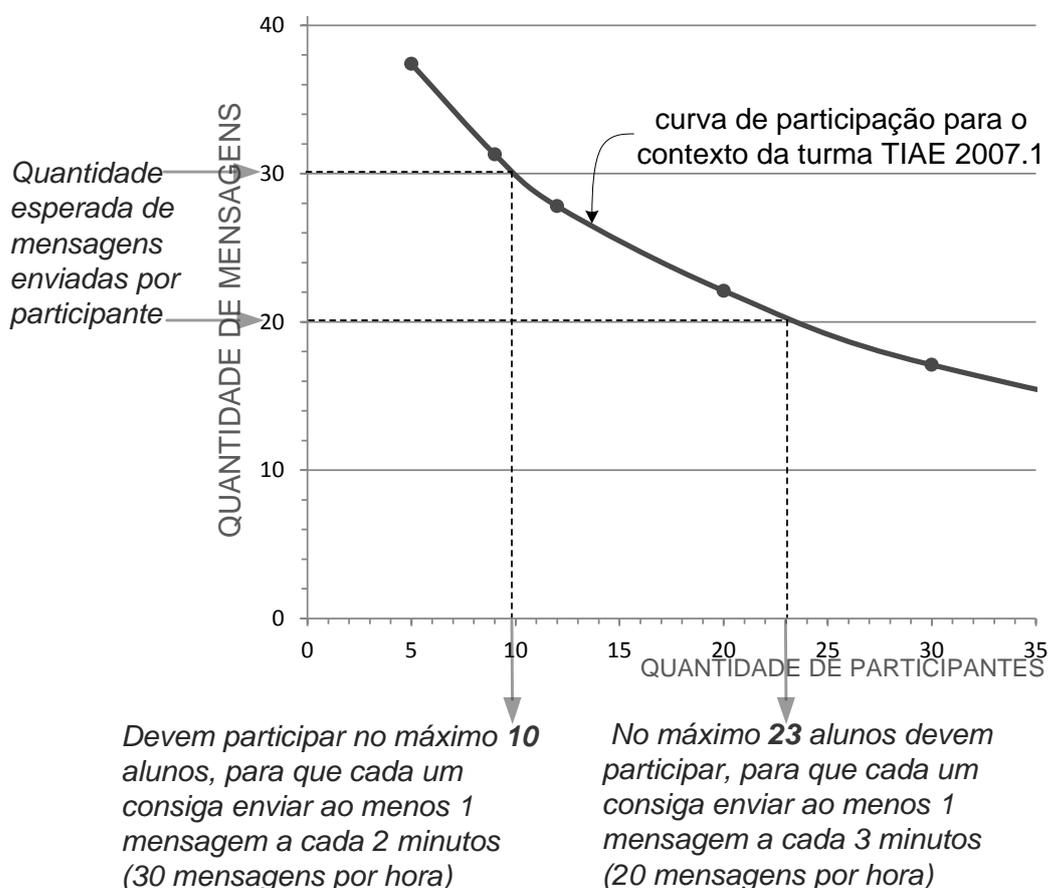


Figura 13. Quantidade de mensagens em função da quantidade de participantes para o contexto da turma TIAE 2007.1

Por exemplo, considerando-se o contexto da turma da Figura 13 ($\text{taxaLeitura} = 12,6 \text{ msg/min}$; $\text{taxaEscrita} = 2,3 \text{ msg/min}$; $\text{taxaIniciativaParaEscrever} = 1,3$), caso o professor espere que cada participante produza ao menos 20 mensagens ao longo da sessão de 60 minutos de bate-papo, então conclui-se que devem participar no máximo 23 alunos. Se o professor espera que cada participante produza 30 mensagens na sessão (1 mensagem a cada 2 minutos), precisa então organizar a turma em grupos ainda menores, e pelo modelo identifica-se que o grupo deveria ter 10 ou menos participantes considerando o perfil médio dos alunos daquela turma.

7. Conclusão: Contribuições da Pesquisa

Na presente pesquisa, quanto à Ciência do Design, foi elaborado um artefato, que é um modelo matemático para representar o comportamento de usuários em sistemas de bate-papo no contexto educacional, visando resolver o problema de determinar qual o número máximo de alunos que devem participar de uma sessão de bate-papo. A partir das estimativas de produção de mensagens por participante em função do tamanho do grupo, foi possível traçar a curva de participação para um determinado contexto de turma e assim visualizar o limite de participantes onde um determinado nível mínimo de participação é garantido.

Os resultados produzidos pelo modelo elaborado mostraram-se adequados em comparação com os dados reais obtidos de 142 sessões de bate-papo realizadas em 18 turmas de uma disciplina a distância, e assim conclui-se que o modelo elaborado nessa pesquisa produz estimativas adequadas. O modelo é útil para subsidiar diferentes estudos, e nessa pesquisa mostrou-se a utilidade do modelo para apoiar o planejamento do tamanho do grupo de alunos que devem participar de uma sessão de bate-papo educacional – esse resultado é de interesse para as instituições que promovem educação *online*.

Quanto à Ciência Comportamental, a presente pesquisa contribuiu para um maior entendimento sobre a teorização da participação em bate-papo educacional. Buscou-se caracterizar o comportamento de participação em função de processos elementares: leitura e escrita de mensagens, e de reflexão, e assim, equacionar o limite máximo de produção de mensagens em função da quantidade de participantes e, a partir do modelo elaborado, foi possível estimar a quantidade esperada dessa produção de mensagens num dado contexto. Com isso, evidenciou-se a influência do aumento do tamanho do grupo na participação dos alunos. Quanto maior a quantidade de participantes, menor é a produção de mensagens, uma vez que o participante necessita de mais tempo para ler as mensagens dos outros participantes.

Uma das limitações da pesquisa foi o pouco estudo realizado sobre os processos elementares de participação: escrita e leitura de mensagens, e reflexão. Em trabalhos futuros, pretende-se caracterizar melhor esses processos. Para o processo de escrita, foi desenvolvido um mecanismo de rastreamento de teclas pressionadas durante uma sessão de bate-papo. Esse mecanismo encontra-se implantado em um bate-papo típico do Portal Tagarelas [Estruc e Pimentel, 2012]. Com a divulgação e uso do Portal Tagarelas, será possível caracterizar melhor o processo de escrita de mensagens em diferentes contextos educacionais. Para os processos de leitura e reflexão, pretende-se realizar estudos com sistemas de rastreamento de retina (*eyetracking*).

Outra limitação da presente pesquisa foi a concentração dos dados usados para a avaliação do modelo estarem ao redor de 10 alunos por sessão, com desvio padrão de 3,8 alunos. Essa concentração impossibilitou a avaliação do modelo em grupos pequenos, com 5 ou menos alunos, e também em grupos maiores, com 15 ou mais alunos. Em trabalhos futuros, espera-se avaliar o modelo considerando uma maior variação no número de participantes em um mesmo contexto.

Também espera-se, em trabalhos futuros, que sejam produzidas estimativas em outros contextos, variando-se as disciplinas e os cursos, visando obter uma estimativa genérica, ainda que grosseira, para as diferentes turmas de diferentes cursos de nível superior, como Pedagogia por exemplo.

Referências

- Aaron, P. G., Joshi, R. M., Gooden, R., & Bentum, K. E. (2008). Diagnosis and treatment of reading disabilities based on the component model of reading an alternative to the discrepancy model of LD. *Journal of Learning Disabilities*, 41(1), 67-84.
- Böhlke, O. (2003). A comparison of student participation levels by group size and language stages during chatroom and face-to-face discussions in German. *CALICO Journal*, 21 (1), 67-87.
- Censo, E. A. D. BR 2012 (2013). Organização Associação Brasileira de Educação a Distância. Disponível em: http://www.abed.org.br/censoead/censoEAD.BR_2012_pt.pdf. Acessado em 29 nov 2014.
- De Souza e Silva, E., Figueiredo, D., Leão, R. (2009). The TangramII integrated modeling environment for computer systems and networks. *SIGMETRICS Perform*.
- Estruc, M., & Pimentel, M. (2012). Portal Tagarelas: bate-papo para educação. In *Anais do Simpósio Brasileiro de Informática na Educação (Vol. 23, No. 1)*.
- Fogliatti, M. C. e N. M. C. Mattos (2007) *Teoria de Filas*. Interciência, Rio de Janeiro.
- Fuks, H., Pimentel, M., Lucena, C.J.P. (2006) R-U-Typing-2-Me? Evolving a chat tool to increase understanding in learning activities. *International Journal of Computer-Supported Collaborative Learning*, v. 1, p. 117-142.
- Graesser, A. C., Hoffman, N. L., & Clark, L. F. (1980). Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior*, 19(2), 135-151.
- Gaines, R. S., Lisowski, W., Press, S. J., & Shapiro, N. (1980). Authentication by keystroke timing: Some preliminary results (No. RAND-R-2526-NSF). RAND CORP SANTA MONICA CA.
- Hevner, A., March, S., Park, J., and Ram, S. (2004) "Design Science in Information Systems Research," *MIS Quarterly* (28:1), 2004, pp. 75-105.
- Hevner, A. R. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, v.19, n. 2, p. 87-92, 2007.
- Hrastinski, S. (2007). Participating in synchronous online education. (Doctoral dissertation, Lund University).
- Hrastinski, S. (2008a). The potential of synchronous communication to enhance participation in online discussions: A case study of two e-learning courses. *Information & Management*, 45, 499-506.
- Hrastinski, S. (2008b). What is online learner participation? A literature review. *Computers & Education*.
- INEP (2013). Censo da Educação Superior 2013. MEC, 2013. Disponível em http://download.inep.gov.br/educacao_superior/censo_superior/apresentacao/2014/colativa_censo_superior_2013.pdf. Acessado em 29 nov 2014.
- Kleinrock, L., (1975), *Queueing Systems - Theory*, Vol. I. Wiley-Interscience.
- Löber, A., Schwabe, G., Grimm, S (2007). Audio vs. chat: The effects of group size on

- media choice. Proceedings of the 40th HICCS Hawaii International Conference on System Sciences.
- Lucena, C.J.P., Fuks, H. (2002). Tecnologias de Informação Aplicadas à Educação (TIAE): Manual do Aprendiz. Monografias em Ciência da Computação n.07/02, PUC-Rio, 2002.
- McKay, J. C., Marshall, P. H., & Heath, G. (2008). An exploration of the concept of design in information systems. In Information Systems Foundations: Answering the Unanswered Questions about Design Research (p. EJ).
- Pimentel, M. G., Fuks, H. e Lucena, C. J. P. (2003) “Debati, debati... aprendi? Investigações sobre o papel educacional das ferramentas de bate-papo”, IX Workshop sobre Informática na Escola - WIE2003, Agosto.
- Piva, D.J., Pupo, R. Gamez, L., Oliveira, S. (2011). EAD na Prática – Planejamento, métodos e ambientes de educação online. Elsevier Editora Ltda.,2011, v.,p 164-170.
- Rocha, E. B., Pimentel, M., & Diniz, M. C. (2012). Modelo para estimar a quantidade de participantes em bate-papo educacional. V Workshop de Teses e Dissertações em Sistemas de Informação. – WTDSI2012.
- Rocha, E. B. (2013). Modelo de Participação em Bate-papo Educacional. 2013. 123 f. Dissertação (Mestrado em Informática) – Universidade Federal do Estado do Rio de Janeiro.
- Sacks, H., Schegloff, E. E., Jefferson, G. (1974) A simplest systematics for the organization of turn-taking for conversation. Language, 50. 1974. p. 696-735.
- Shimakura, S.E. (2006) Correlação. In: CE003 - Estatística II. Paraná: Dep. de Estatística da Universidade Federal do Paraná. <http://leg.ufpr.br/~silvia/CE003/node74.html>. Acessado em 20 julho 2013.
- Silva, M. Avaliação da Aprendizagem em Educação Online (2011). Entrevista publicada no YouTube: <http://youtu.be/S7uUd6afEYE>. Acessado em 12 ago 2013.
- Stahl, G. Studying Virtual Math Teams (2009). New York, NY: Springer.
- Wieringa, R. J. (2014). Design Science Methodology for Information Systems and Software Engineering. Heidelberg: Springer.