

Submission: 29/Sep/2022;
Camera ready: 02/Mar/2023;

1st round notif.: 27/Nov/2022;
Edition review: 14/Mar/2023;

New version: 13/Dec/2022;
Available online: 26/Mar/2023;

2nd round notif.: 28/Feb/2023;
Published: 26/Mar/2023;

Usabilidade dos Ambientes Virtuais de Aprendizagem Canvas e Blackboard: Caso de estudo em uma Universidade Brasileira

Title: *Usability of Canvas and Blackboard Learning Management Systems: A case study at a Brazilian University*

Lucas Exposto Soares

Universidade Virtual do Estado de São Paulo (Univesp)

ORCID: 0000-0001-8772-6899

lucas.exposto@gmail.com

José Maria Clementino Júnior

Universidade Virtual do Estado de São Paulo (Univesp)

ORCID: 0000-0001-8281-4152

juninhoclementino45@gmail.com

Lucas Augusto Vieira Brito

Universidade Virtual do Estado de São Paulo (Univesp)

ORCID: 0009-0007-5990-8442

lucasaugusto.vb@gmail.com

Rogério Gil Gonzaga

Universidade Virtual do Estado de São Paulo (Univesp)

ORCID: 0009-0002-3130-6128

rogeriogilgonzaga@gmail.com

Humberto Antonio Antico

Universidade Virtual do Estado de São Paulo (Univesp)

ORCID: 0000-0001-7253-5168

humberto.antico@gmail.com

Rodolfo Jardim de Azevedo

Universidade Virtual do Estado de São Paulo (Univesp)

Universidade Estadual de Campinas (UNICAMP)

ORCID: 0000-0002-8803-0401

rodolfo.azevedo@univesp.br

Alessandra Alaniz Macedo

Universidade Virtual do Estado de São Paulo (Univesp)

Universidade de São Paulo (USP)

ORCID: 0000-0001-5271-3086

ale.alaniz@usp.br

Resumo

Para amparar o processo de ensino-aprendizagem, as instituições de ensino devem avaliar cuidadosamente a usabilidade de seus Ambientes Virtuais de Aprendizagem (Learning Management Systems - LMS), antes de entregá-los a professores e alunos. Este artigo apresenta o processo e os resultados das avaliações de usabilidade em atividades de ensino-aprendizagem apoiadas pelos LMS Canvas e Blackboard em uma universidade brasileira. A metodologia considerou testes com usuários para avaliar cinco atributos de usabilidade nos sistemas. Além disso, especialistas em usabilidade realizaram inspeções de usabilidade para analisar problemas e possíveis soluções. Os resultados indicaram que ambas as plataformas apresentam graves problemas de usabilidade. Em termos de comparações entre os dois sistemas LMS, os resultados não mostraram uma disparidade significativa de usabilidade entre eles. Como resultado, também apontamos os principais problemas e soluções dos LMS.

Palavras-chave: Interface Humano Computador; Usabilidade; Intereração; Inspeção; Ambientes Virtuais Aprendizagem; EaD; Canvas; Blackboard.

Abstract

Educational institutions should carefully evaluate the usability of their Learning Management Systems (LMS) before delivering them to teachers and students. This paper presents the process and the results of the usability evaluations

Cite as: Soares, L. E., Antico, H. A., Clementino Júnior, J. M., Azevedo, R. J., Brito, L. A. V. & Macedo, A. A.. (2023). Usabilidade dos Ambientes Virtuais de Aprendizagem Canvas e Blackboard: Caso de estudo em uma Universidade Brasileira. *Revista Brasileira de Informática na Educação*, v. 31, 149-173. DOI: 10.5753/rbie.2023.2921.

in the teaching-learning activities supported by the Canvas and the Blackboard LMS at a Brazilian University. The methodology considered tests with users to evaluate five usability features in the Canvas and Blackboard systems. In addition, usability experts performed usability inspections to analyze problems and possible solutions. The results indicated that both platforms have severe usability issues. In terms of the comparisons between the two LMS systems, the results did not show a significant disparity in usability between them. As a result, we also pointed out the main problems and solutions of the LMS.

Keywords: *Human Computer-Interface; Usability; Interaction; Inspection; Learning Management System; EaD; Canvas; Blackboard.*

1 Introdução

Os avanços ocasionados nas últimas décadas devido à evolução da Internet e suas tecnologias impulsionaram os métodos de difusão e de obtenção do conhecimento. Atualmente, é perceptível o crescimento da Educação a Distância (EaD), seja em uma evolução natural ou resultante da necessidade emergencial devido à presente situação global da pandemia do SARS-CoV-2/Covid-19 (Sobaih, Hasanein, & Abu Elnasr, 2020; Gumasing, Vasquez, Doctora, & Perez, 2022). O aumento da adoção de EaD também está sendo impulsionado e solidificado pelo suporte de plataformas do tipo *Learning Management System* (LMS), também conhecidas como Ambientes Virtuais de Aprendizagem (AVA) (Oliveira, Procaci, & Siqueira, 2020). Os sistemas LMS são amplamente utilizados para criar, administrar e disponibilizar conteúdo e materiais de aprendizagem, pretendendo aumentar a eficiência do ensino e enriquecer o aprendizado dos alunos (Chen, Sanderson, Nichshyk, Bong, & Kessel, 2021).

Com o objetivo de atender à crescente demanda por EaD, uma *Universidade Brasileira* com milhares de alunos nesta modalidade de ensino, vem buscando se manter atualizada em termos de plataformas digitais robustas e eficientes no suporte do processo de ensino-aprendizagem. Recentemente, a instituição realizou a migração da plataforma LMS Canvas para a LMS Blackboard. O LMS Canvas utilizado se tratava de uma versão de código aberto da plataforma instalada localmente e na nuvem da própria instituição. Os custos desse LMS eram decorrentes da infraestrutura de nuvem e da equipe de suporte disponível 24 horas por 7 dias da semana. A demanda por pessoal dedicado chamava atenção para questões de escalabilidade do LMS, devido ao receio de indisponibilidade de serviços técnicos e de atividades educacionais gerado por falta de pessoal específico para determinados momentos (atualizações de versões, entregas de trabalhos e revisões das disciplinas e configurações). Infelizmente, esses riscos geraram alterações nos processos educacionais disponibilizados pelo LMS.

A troca de LMS, motivada pelas questões de custo e pela oportunidade de mudança, foi realizada por meio de um processo de licitação com um Termo de Referência considerando escalabilidade, *service-level agreement* alto e garantia de migração de dados. Participaram três concorrentes na licitação e a plataforma LMS Blackboard foi a ganhadora. O fornecedor do LMS Canvas não participou, mesmo que ciente da licitação.

Substituições de plataformas de ensino, assim como trocas de sistemas gerenciais em empresas, ocasionam certa resistência e adaptação em sua utilização por seus usuários. Não obstante, oferecer uma plataforma on-line atrativa, eficiente, acessível, imersiva e segura para o usuário exige planejamento e avaliação de uso (Gunesekera, Bao, & Kibelloh, 2019). Nesse sentido, existem maneiras de avaliar se um determinado produto educacional satisfaz quesitos de eficiência e de satisfação do usuário como os testes e as inspeções de usabilidade (Nielsen & Loranger, 2006). Para a Organização Internacional de Padronização, usabilidade é definida como a medida em que um produto pode ser usado por usuários específicos para atingir metas específicas com: eficácia, eficiência e satisfação em um contexto de uso especificado (Associação Brasileira de Normas Técnicas, 2002).

Motivada por esse contexto, a proposta principal do artigo é apresentar a avaliação de usabilidade dos ambientes virtuais de aprendizagem Canvas e Blackboard no processo de ensino-aprendizagem de uma *Universidade Brasileira*, a fim de realizar uma comparação entre os LMS

para tentar entender as perdas e os ganhos da migração na instituição. Para esse objetivo, aplicou-se testes de usabilidade e inspeções de usabilidade. A aplicação de testes de usabilidade¹ permitiu que alunos e professores pudessem expressar a experiência de uso em cada LMS. Os dados coletados foram analisados para descrever a percepção do usuário-final. A proposta também contou com inspeções² realizadas por especialistas de IHC, durante o uso dos LMS. Essas inspeções foram guiadas pelas heurísticas de Nielsen. Os resultados deste trabalho contribuem com a instituição em questão, mas também com outras instituições que usam ou desejam usar esses LMS ou similares. Espera-se também que este trabalho seja um exemplo para (i) guiar outros pesquisadores a compararem sistemas; e (ii) contribuir no entendimento do impacto dos LMS, principalmente na EaD e no cenário da educação de nível superior no Brasil.

2 Trabalhos Relacionados

Plataformas LMS são fundamentais para disponibilizar conteúdos de aulas, atividades, exercícios e avaliações, além de materiais complementares, para os alunos. Nesta categoria de *software*, há diversas opções de sistema com diferentes funcionalidades e características de interação via suas interfaces de usuário. Portanto, é indispensável um estudo da literatura, a fim de mapear os trabalhos que avaliaram o quesito usabilidade de LMS em outras universidades, principalmente dos sistemas Canvas e Blackboard. Assim, a seguir são descritos trabalhos que elencam os LMS citados e aplicam diferentes abordagens para mapeamento e comparação das interfaces de usuários envolvidas, de modo que contribuem e são pertinentes à pesquisa aqui apresentada.

Nesse contexto, Alturki, Aldraiweesh, & Kinshuk (2016) avaliaram a usabilidade e a acessibilidade do Blackboard na King Saud University, buscando responder questões relacionadas aos recursos essenciais, de navegação, de eficácia de uso, de custo de aquisição e de barreiras de uso da plataforma. Neste estudo, foi selecionada uma amostra de 400 pessoas para execução de um treinamento com diferentes tarefas no sistema e posterior avaliação da plataforma, por meio de um questionário. Os elementos de usabilidade e acessibilidade avaliados incluíam o *design* da interface, o uso interativo dos elementos da interface e a funcionalidade dos elementos de navegação para acessar materiais de instrução. Os resultados mostraram uma classificação média de 4 em relação à acessibilidade e usabilidade, considerando uma escala de 1 a 5. Além disso, mais de 56% dos entrevistados concordaram que o Blackboard tinha facilidade de uso, bom *design* e funcionalidades altamente interativas, com todos os recursos necessários. No entanto, os membros da universidade avaliaram negativamente o *software* quanto à sua usabilidade para acessar testes e questionários.

Em um estudo de 2018, Grossi, de Souza Elias, Chamon, & Leal (2018) compararam parâmetros tecnológicos dos AVAs Moodle e Canvas, a fim de facilitar a escolha de um ambiente virtual a ser utilizado para finalidade educacional. Nesta investigação, um dos parâmetros tecnológicos avaliados foi a usabilidade. Os aspectos selecionados permitiram avaliar a usabilidade dos

¹É uma técnica de pesquisa para avaliar um produto ou serviço, envolvendo usuários representativos do público-alvo (experientes ou não) na realização de tarefas típicas, enquanto um analista observa, ouve e anota (Rocha & Baranauskas, 2003)

²São métodos de avaliação de *design* cujos avaliadores (especialistas, consultores ou usuários finais) inspecionam ou examinam os aspectos relacionados à usabilidade (Rocha & Baranauskas, 2003)

sistemas, usando como referência a ISO 9241 e as heurísticas de Nielsen. Os resultados indicaram que o Canvas foi mais completo em termos de funcionalidades, colaborativo e com interface de usuário muito intuitiva. Apesar de o Canvas ter obtido uma classificação melhor do que o Moodle devido à navegação mais intuitiva do usuário, os dois AVAs apresentaram interfaces claras e de fácil compreensão, segundo os autores.

Em Ahmad, Beyene, & Giannoumis (2018), os autores avaliaram a capacidade do usuário em aprender e analisar acessibilidade em duas plataformas diferentes, o Fronter e o Canvas. Em relação à usabilidade, foram designadas tarefas aos usuários que nunca tiveram contato com ambas as plataformas com intuito de avaliar a capacidade do usuário aprender sozinho. Baseado no método de entrevista, os participantes expressavam suas opiniões durante a execução das tarefas. Ao final, a plataforma Canvas foi classificada como a melhor opção por oferecer uma melhor usabilidade de modo geral, perdendo para o Fronter em situações relacionadas ao *design* visual, como esquema de cores e *design* gráfico.

Em Alhadreti (2020), os autores compararam experimentalmente a usabilidade do Blackboard (BB) e do Desire2Learn (D2L), usados na Universidade Umm Al-Qura, no Reino da Arábia Saudita. A comparação envolveu os seguintes critérios: a quantidade e a qualidade dos problemas de usabilidade descobertos, a eficácia dos sistemas, bem como a eficiência deles e a satisfação dos usuários. Um conjunto de tarefas foi elaborado para cada LMS e posteriormente aplicado no LMS com um grupo de 20 alunos. O método utilizado para avaliar as ferramentas foi o *think-aloud*³. Os resultados estatísticos indicaram que o Blackboard tem diversos problemas graves de usabilidade, uma vez que a conclusão da tarefa foi mais lenta no BB do que no D2L, exigindo mais cliques do mouse e mais páginas visitadas. Além disso, os participantes também cometeram mais erros no BB. No entanto, os níveis de satisfação do usuário foram os mesmos para os dois sistemas.

Andrews-Todd, Bond, Speller, Stevenson, & Tillman (2013) realizaram uma pesquisa com o objetivo de investigar como a usabilidade do Blackboard impactou as experiências de ensino de alunos e professores na *Historically Black College/University* (HBCU), a partir de percepções cognitivas, organizacionais e socioambientais. Eles utilizaram métodos de extração de protocolos verbais para posteriormente realizar a interpretação das atividades. Essa extração buscou o significado de como os alunos percebem o uso do sistema pelo instrutor e como os instrutores usam o sistema para beneficiar o aprendizado do aluno. Uma das constatações foi que a interface do Blackboard é de fácil utilização por alunos. Entretanto, também foi percebida a frustração do instrutor com a quantidade de tempo necessária para configurar e manter recursos no Blackboard, como o tempo gasto inserindo notas e configurando avaliações. O objeto específico da atividade era associar sentimentos negativos em relação ao uso dos recursos notas e avaliação no LMS. Os autores observaram que os problemas de usabilidade do BB incluem o envio de materiais por meio do *Dropbox* digital e funções de avaliação. Um benefício importante para os alunos foi descoberto no estudo: um LMS aumenta a memória prospectiva para eventos futuros relacionados ao curso. Outra importante observação foi de que informações não postadas na plataforma (principalmente notas), de acordo com o esperado pelos alunos, gera uma opinião negativa sobre sua experiência de ensino.

O estudo de Gumasing et al. (2022) teve como objetivo determinar a pontuação de aceitabi-

³*Think-aloud* é um método usado para coletar dados, o qual consiste em instigar os participantes a falarem, enquanto realizam um conjunto de tarefas específicas (Alhadreti, 2020).

dade e comparar a usabilidade do Blackboard e do Canvas, com base na satisfação dos alunos. Os pesquisadores utilizaram um questionário de escala de usabilidade do sistema para avaliar os sistemas com base na pontuação de aceitabilidade calculada, a partir dos dados coletados. O estudo também utilizou a estatística descritiva e o teste *t* para tentar identificar uma diferença significativa entre o Blackboard e o Canvas. Os autores concluíram que não existe diferença significativa entre a usabilidade e a satisfação do usuário nos dois sistemas.

O artigo de Chen et al. (2021) buscou desafios de usabilidade para instrutores no sistema Canvas. Para isso, seis tarefas foram atribuídas a 35 instrutores universitários em disciplinas de informática e de engenharia. Os participantes foram instruídos a expressar suas observações positivas ou negativas durante a execução das tarefas, as quais foram coletadas e analisadas juntamente com dados de entrevistas realizadas, antes e após, a conclusão das tarefas. Os resultados mostraram que, embora o Canvas tenha feito esforços contínuos para melhorar a sua usabilidade, os instrutores ainda enfrentam alguns desafios de usabilidade. A maioria desses problemas estavam relacionados às expectativas de onde encontrar ou como usar as ferramentas.

Os trabalhos apresentados contribuíram direta ou indiretamente para a realização da proposta deste artigo, por meio do aprendizado de técnicas e de métodos como o levantamento de dados, por meio de questionário ou formulários. A comparação dos trabalhos destaca que, em Alturki et al. (2016) e Gumasing et al. (2022), os resultados foram encontrados a partir de análises quantitativas dos dados. Em Ahmad et al. (2018), Alhadreti (2020), Andrews-Todd et al. (2013), Grossi et al. (2018) e Chen et al. (2021), os resultados foram caracterizados por análises qualitativas, que trazem estratégias de melhorias para as plataformas estudadas, ou em apontamentos de recursos que são eficientes ou não.

Em Alturki et al. (2016), Andrews-Todd et al. (2013) e Chen et al. (2021), apenas um sistema LMS é analisado, buscando identificar características quanto à usabilidade, aplicabilidade e apontamento de recursos mais eficientes. Em Grossi et al. (2018), Ahmad et al. (2018), Alhadreti (2020) e Gumasing et al. (2022), os trabalhos também identificam tais características, porém, com a diferença de realizarem a análise comparativa entre dois sistemas LMS. De qualquer forma, estes trabalhos possuem aspectos comuns na realização de análises de resultados, em relação aos níveis de satisfação de usuários finais, sejam alunos ou professores.

A metodologia aplicada aqui difere dos trabalhos apresentados em alguns pontos, uma vez que nenhum dos estudos avaliou os resultados de usabilidade do ponto de vista de usuários finais (alunos e professores) e de especialistas da computação. Além disso, é importante destacar que não foram encontradas pesquisas que vislumbram um estudo comparativo entre os sistemas LMS Canvas e Blackboard considerando alunos e professores.

3 Materiais e Métodos

Este trabalho foi realizado a partir da aplicação dos seguintes métodos para avaliar a usabilidade dos sistemas LMS Canvas e Blackboard, dado que esses métodos podem ser realizados em um curto período e com custos satisfatórios para o projeto: 1) testes de usabilidade, considerando a interação com a plataforma avaliada e a posterior aplicação de questionários de satisfação realizados pelos usuários; 2) inspeção realizada por especialistas. Todos os procedimentos deste artigo

foram conduzidos seguindo as normas e diretrizes de ética em pesquisas.

A Figura 1 apresenta o fluxograma da metodologia para avaliação de usabilidade dos sistemas Canvas e Blackboard. Os métodos investigados e as atividades associadas estão detalhados nas seções a seguir.

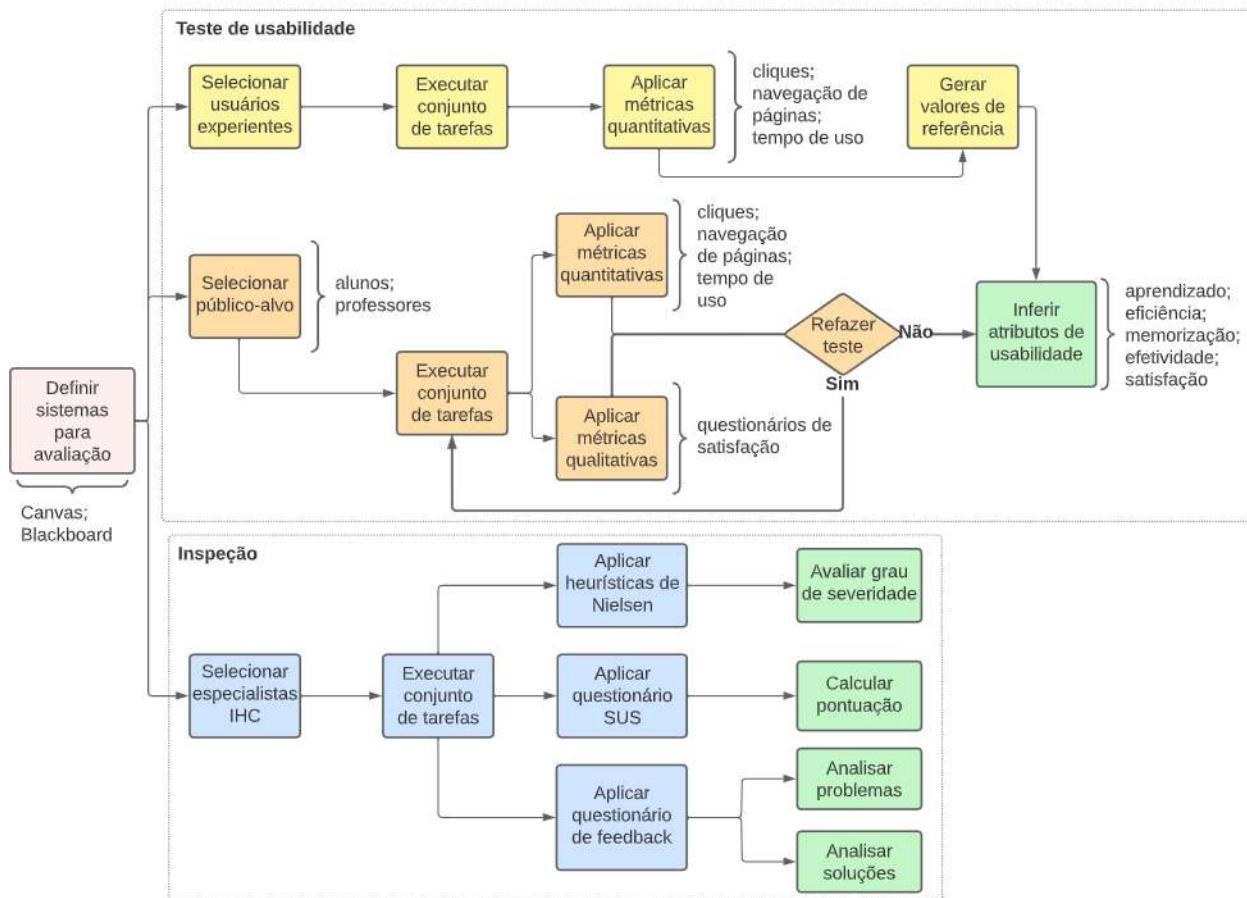


Figura 1: Fluxograma da metodologia para avaliação de usabilidade.

3.1 Testes de Usabilidade

A avaliação de usabilidade, a partir de testes com usuários, foi o primeiro método executado neste estudo. Para o desenvolvimento desta avaliação, utilizou-se a metodologia adaptada de Phongphaew & Jiamsanguanwong (2018) com as seguintes etapas: (A) Definições iniciais, (B) Aplicação do teste e (C) Validação, ilustradas na Figura 2. Essas etapas foram adaptadas para serem aplicadas em uma versão reduzida da metodologia em termos de parâmetros, mas ampliada em termos de usuários da seguinte maneira: (i) Escala Likert de cinco pontos em vez de sete (discreto fortemente: 1, e concordo fortemente: 5), (ii) cinco tarefas para alunos e professores em vez de sete para alunos e doze para professores da metodologia original, (iii) duas questões no Questionário Pós-Cenário (QPC) ao invés de três, e treze questões no Questionário de Usabilidade do Sistema Pós-Etapa (QUSPE) ao invés de dezenove, acompanhando o formato reduzido, (iv) dez estudantes e dez professores, diferentemente da proposta que utilizou cinco e cinco, respectivamente.

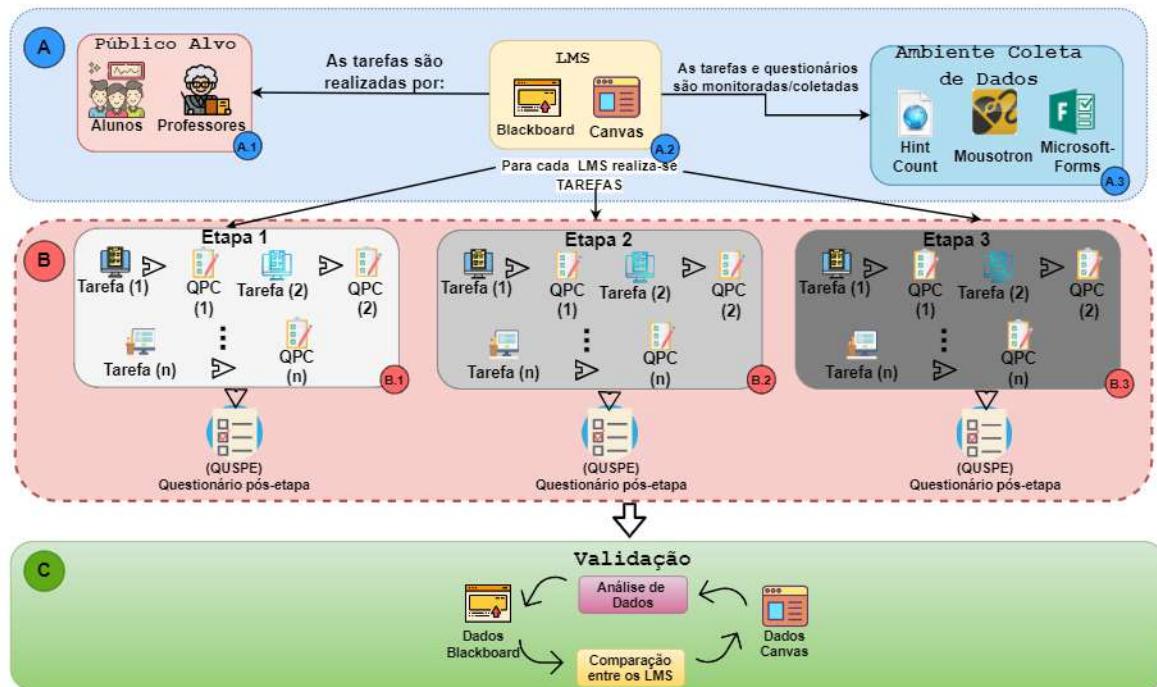


Figura 2: Visão geral dos testes de usabilidade com usuários com a metodologia adaptada de Phongphaew & Jiamsanguanwong (2018). Têm-se as seguintes etapas (A) Definições iniciais, (B) Aplicação do teste e (C) Validação.

Para a compilação dos resultados, utilizou-se a NBR ISO 9241-11 (Associação Brasileira de Normas Técnicas, 2002) devido à sua abrangência e os atributos de usabilidade propostos por Nielsen & Loranger (2006) nas seguintes métricas tradicionais para essa medida:

- (a) *Capacidade de aprender a utilizar o sistema*: refere-se à capacidade dos participantes conseguirem realizar as tarefas sem conhecimento prévio da plataforma, conseguindo navegar sobre diferentes funcionalidades. Este atributo foi representado pelo tempo de finalização de cada tarefa na primeira vez que usam o sistema. Um menor tempo significa facilidade em aprender a usar o sistema.
- (b) *Eficiência*: refere-se aos participantes realizarem as tarefas em um menor tempo possível de maneira correta, após estarem familiarizados com o sistema. Dessa forma, um tempo menor indica uma melhor eficiência do sistema.
- (c) *Capacidade de memorização*: refere-se à capacidade dos participantes em memorizar o caminho da navegação para concluir a tarefa. Portanto, é utilizado o tempo de conclusão de cada tarefa após não utilizar o sistema por 7 dias. Uma duração menor significa uma melhor memorização.
- (d) *Efetividade*: refere-se à efetividade de execução da tarefa pelos participantes na primeira etapa, de forma a alcançar o objetivo. O atributo de efetividade é definido como a acurácia e a completnude, o qual é semelhante ao atributo de erro de Nielsen. O Índice de Sucesso da Tarefa (IST) é o produto da taxa de conclusão multiplicada pela taxa de acurácia. A taxa de conclusão é o número real de páginas navegadas dividido pelo número esperado de trocas de páginas. A taxa de acurácia é o número real de cliques dividido pelo número esperado de cliques. Valores de IST iguais ou próximos de 1 indicam melhor efetividade na tarefa.
- (e) *Satisfação*: refere-se ao nível de satisfação dos participantes, após o contato com o sistema LMS. Neste estudo, o grau de satisfação foi coletado por meio de questionários de pontuação após a conclusão da tarefa e da etapa.

Os testes foram aplicados com alunos e professores, que realizaram cinco tarefas rotineiras, cujos recursos estão comumente disponibilizados em ambas as ferramentas LMS: Canvas e Blackboard. Paralelamente à execução das tarefas, utilizou-se um formulário para capturar os níveis de satisfação geral de utilização das ferramentas, bem como os níveis de dificuldade e agilidade para a realização de cada tarefa.

Além da satisfação do usuário, por meio de medições objetivas como tempo, navegação de página e cliques do usuário ao executar as tarefas, buscou-se inferir a capacidade de aprendizado e da memorização dos usuários nas tarefas executadas, bem como, a eficiência e efetividade das ferramentas.

Os detalhes de cada etapa da metodologia de avaliação de usabilidade deste trabalho são apresentados a seguir.

3.1.1 Definições Iniciais

Os primeiros passos do teste de usabilidade foram a escolha (A.1) do público-alvo, (A.2) dos sistemas LMS e (A.3) do ambiente de coleta dos dados, conforme a Figura 2.

Em relação ao **público-alvo** (Figura 2-A.1), para evitar qualquer tipo de viés na pesquisa, o primeiro critério utilizado para a seleção foi convidar somente participantes que nunca tiveram qualquer tipo de vínculo com a instituição-alvo. O segundo critério definido foi selecionar somente participantes que não possuíam experiência de uso nos sistemas avaliados. O terceiro critério foi selecionar somente participantes que tinham vínculo de aluno ou de professor em alguma universidade brasileira pública, ou particular. Finalmente, a amostra utilizada deveria representar variedades de categoria, idade, gênero, área e nível de habilidade referente à utilização de computadores e sistemas LMS.

Dessa maneira, foram selecionados dez estudantes (sete alunas e três alunos com média de idade de 36 anos) e dez professores (três professoras e sete professores com média de idade de 42 anos). De acordo com experimentos e constatações de Nielsen (2012), um teste de usabilidade com vinte usuários é suficiente para obtenção de resultados estatisticamente significativos, sendo capaz de revelar ao menos 95% dos problemas de usabilidade (Faulkner, 2003). Destaca-se que dentre o público-alvo de vinte usuários no total, o nível de habilidade referente à utilização de computadores era variado. De maneira semelhante, foram observados variados graus de aptidão no uso de sistemas LMS entre os participantes, sendo que 50% dos professores tinham mais de cinco anos de experiência e 70% dos alunos possuíam entre seis meses e um ano de experiência. Os participantes eram das diferentes áreas do conhecimento.

Dentre os **sistemas LMS** existentes, as plataformas Canvas e Blackboard foram selecionadas para esta investigação, dado que a instituição de ensino, que é parte desta pesquisa, estava em um processo de migração e fazia o uso de ambos os sistemas (Figura 2-A.2), durante o desenvolvimento da pesquisa.

O **ambiente de coleta dos dados** diz respeito às ferramentas utilizadas durante o processo de coleta de dados, monitoramento das tarefas e armazenamento das questões na pesquisa (Figura 2-A.3). Para a coleta de informações sobre navegações entre páginas, cliques e tempo, foram

utilizados os softwares *Hit Count*⁴ e *Mousotron*⁵. Para criação e coleta dos questionários foi utilizado o *Google Forms*⁶.

3.1.2 Aplicação do Teste

Em um primeiro contato, os participantes foram informados do contexto da pesquisa. Em seguida, aplicou-se um questionário de triagem, a fim de coletar as informações pessoais e demográficas de cada participante. É importante ressaltar que, apesar de serem avaliados os dois sistemas LMS (Canvas e Blackboard) com alunos e com professores, as tarefas, questionários e avaliação seguiam a mesma estrutura.

Conforme ilustra a Figura 2 (B), um **conjunto de tarefas** foi atribuído para os dois grupos de usuários (alunos e professores) que não tinham vínculo com a Universidade, a fim de evitar qualquer viés na pesquisa. Durante a execução das tarefas, **métricas quantitativas** (ex. tempo de uso e cliques) e **qualitativas** (ex. satisfação do usuário) foram utilizadas para avaliar a usabilidade dos sistemas. A aplicação do teste se deu por **etapas**, as quais estão descritas a seguir.

Para execução do Canvas e do Blackboard, foi criado um **Conjunto de Tarefas** com base na existência de recursos funcionais e de particularidades em relação aos acessos, navegação e o *design* das telas de cada LMS para suportá-lo. Por isso, os graus de facilidade e de agilidade que os usuários realizaram as tarefas nas duas ferramentas foram aspectos essenciais para obtenção dos resultados de análise comparativa entre as ferramentas.

Todos os dez alunos que participaram da pesquisa realizaram as seguintes cinco tarefas no Canvas e no Blackboard para a 3^a semana da disciplina de “Pensamento Computacional”: (i) responder à atividade avaliativa da disciplina, (ii) fazer *download* do material de *slides* de apoio de um vídeo, na 3^a semana da disciplina, (iii) adicionar um novo comentário com o texto “teste” no tópico do fórum de dúvidas, (iv) visualizar as notas da disciplina e (v) enviar mensagem com o texto “teste” ao “Professor Teste01” da disciplina. É interessante ressaltar que a disciplina de Pensamento Computacional foi escolhida por ser uma disciplina de formação básica para as diferentes formações profissionais da universidade.

Para os dez professores, foram determinadas outras cinco tarefas similares às dos alunos para execução no Canvas e no Blackboard para a 3^a semana da disciplina de “Pensamento Computacional”. As tarefas foram: (i) enviar mensagem com o texto “teste” para o “Aluno Teste01” da disciplina, (ii) fazer *download* do material de *slides* de apoio de um vídeo, na 3^a semana da disciplina, (iii) criar um tópico de discussão “teste” no fórum de dúvidas da semana na disciplina, (iv) visualizar o quadro geral (atividades e avaliações) de nota de todos os alunos da disciplina e (v) visualizar interação dos alunos na disciplina.

Para definir os tipos de **Medições Quantitativas** dos usuários durante a realização das tarefas, utilizou-se o software *Mousotron* para capturar a quantidade de cliques e o tempo necessário para realização de cada tarefa. Também foi utilizado o software de complemento *Hit Count* disponível para o navegador Google Chrome. Esse recurso teve a função de auxiliar na captura da quantidade de páginas navegadas, quando cada tarefa é realizada. A fim de estabelecer um valor

⁴[https://chrome.google.com/webstore/detail hit-count/dofbfldpbmpmljhjchkijcekdjlalojg](https://chrome.google.com/webstore/detail	hit-count/dofbfldpbmpmljhjchkijcekdjlalojg)

⁵<https://www.blacksunsoftware.com/mousotron.html>

⁶<https://www.google.com/intl/pt-BR/forms/about/>

referencial de tempo, trocas de páginas e cliques necessários para execução de cada tarefa, três usuários externos ao teste e experientes realizaram as tarefas propostas para alunos e professores. A média desses valores, chamado de **referência**, foi calculada para posterior comparação com os valores alcançados pelos usuários participantes do teste.

As **Medições Qualitativas** ocorreram com a aplicação de questionários de satisfação do usuário. Para essa medição, utilizou-se o *Microsoft Forms*, uma ferramenta online para elaboração de Survey. Os seguintes dois tipos de questionários foram elaborados:

(a) *Questionário Pós-Cenário (QPC)*: Ao final de cada uma das cinco tarefas realizadas, o usuário respondeu ao QPC de Lewis (1995), o qual é composto da avaliação de duas questões, utilizando a Escala Likert de cinco pontos (discordo fortemente: 1, e concordo fortemente: 5). Para medir a satisfação do usuário, calcula-se a média das duas questões-item, sendo que quanto maior o número, maior a satisfação do usuário.

(b) *Questionário de Usabilidade do Sistema Pós-Etapa (QUSPE)* de Lewis (1995) foi aplicado ao final de cada etapa, sendo composto de treze questões, utilizando a Escala Likert de 5 pontos (discordo fortemente: 1, e concordo fortemente: 5), a fim de avaliar a satisfação geral do usuário em relação à usabilidade do sistema.

O teste compreendia **três etapas subsequentes**, ilustradas por B.1, B.2 e B.3 na Figura 2. Cada etapa era composta de cinco tarefas para execução em cada sistema LMS.

Na etapa B.1 da Figura 2, com o intuito de avaliar a capacidade de aprendizagem e a efetividade nos dois sistemas, os participantes, sem conhecimento prévio sobre as plataformas, realizaram as tarefas designadas. Ao fim de cada tarefa foi aplicado o QPC e, ao final das cinco tarefas, os participantes responderam o QUSPE, marcando assim o fim dessa etapa.

Já a etapa B.2 da Figura 2 foi realizada logo que os usuários terminaram a primeira etapa, sem intervalo de dias. Essa ação objetivava avaliar a eficiência e a efetividade, uma vez que já houve o contato com as tarefas na primeira etapa. Os usuários realizaram as mesmas tarefas da etapa B.1 e foram aplicados os mesmos questionários. Simplificadamente, a etapa B.2 foi uma repetição da etapa B.1. Na etapa B.3 da Figura 2, os participantes foram testados com o objetivo de avaliar a memorização e a eficiência. Assim, esta etapa foi realizada após sete dias da realização da etapa B.2 e, nesse período, os alunos e professores não tiveram acesso aos sistemas LMS. Novamente, os mesmos questionários das etapas B.1 e B.2 foram aplicados.

3.1.3 Validação

Após a realização de todas as etapas para cada LMS, os dados coletados foram submetidos à análise para comparação entre os sistemas estudados (veja Figura 2-C).

3.2 Inspeção com Especialistas

Realizou-se também a inspeção com especialistas nas duas plataformas LMS com o objetivo de apresentar uma análise de usabilidade mais robusta e qualitativa. A Figura 3 ilustra o fluxograma da inspeção realizada, composto por (A) Definições iniciais, (B) Aplicação das tarefas e heurísticas, (C) Aplicação do questionário SUS (*System Usability Scale*), (D) Aplicação do questionário de *feedback* dos especialistas e (E) Validação.

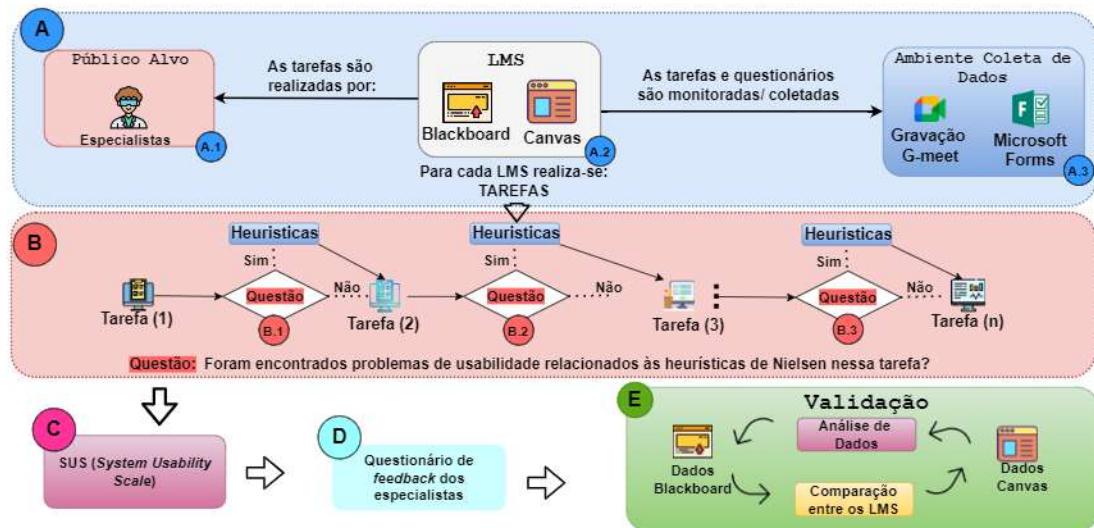


Figura 3: Visão geral do método de inspeção com especialistas.

3.2.1 Definições Iniciais

De acordo com a Figura 3, em relação ao **público-alvo** (A.1), foram escolhidos seis especialistas que tinham experiência em usabilidade, *User Experience* (UX) ou Interação Humano-Computador (IHC). Dos especialistas, foram solicitadas suas experiências na área e constatou-se que 16,7% deles tinham entre 1 e 2 anos, 50% entre 2 e 5 anos e 33,3% possuíam mais de 5 anos de experiência. Novamente, Blackboard e Canvas foram os **sistemas LMS** (A.2) analisados. As plataformas *Google Meet* e *Google Forms* foram utilizadas para realizar a **coleta de dados** (A.3).

3.2.2 Aplicação das tarefas e heurísticas

As seguintes sete tarefas foram designadas para a inspeção dos LMS, na mesma disciplina de Pensamento Computacional do teste de usabilidade, que é de formação básica do currículo da universidade em questão:

- *Tarefa 1*: Fazer o *download* do material de *slides* de apoio do vídeo 1, na semana 3 da disciplina “Pensamento Computacional”;
- *Tarefa 2*: Criar um tópico de discussão “teste” no fórum de dúvidas da semana 3 na disciplina “Pensamento Computacional”;
- *Tarefa 3*: Visualizar o quadro geral (atividades e avaliações) de nota de todos os alunos da disciplina “Pensamento Computacional”;
- *Tarefa 4*: Visualizar a interação dos alunos na disciplina “Pensamento Computacional”;
- *Tarefa 5*: Responder à atividade avaliativa da semana 3 da disciplina “Pensamento Computacional”;
- *Tarefa 6*: Adicionar um novo comentário com o texto “teste” no tópico do fórum de dúvidas da semana 3 na disciplina “Pensamento Computacional”;
- *Tarefa 7*: Enviar mensagem com o texto “teste” ao “Professor Teste01” da disciplina “Pensamento Computacional”.

Conforme apresenta a Figura 3 (B.1, B.2 e B.3), a dinâmica da inspeção consistiu em realizar

as tarefas, uma de cada vez, e após a conclusão de cada tarefa, o participante respondia em um questionário se havia encontrado algum problema de usabilidade. Caso encontrasse, o especialista deveria apontar a(s) **heurística(s) de Nielsen** (Nielsen, 1994) afetada(s), considerando os seus respectivos graus de severidade.

A **Avaliação das Heurísticas de Nielsen** suporta a inspeção sistemática da “interface de usuário” em busca de problemas de usabilidade por especialistas⁷. Esses problemas não atendem as seguintes diretrizes de Nielsen (1994): (i) Visibilidade do estado do sistema, (ii) Compatibilidade entre o sistema e o mundo real, (iii) Controle e liberdade do usuário, (iv) Consistência e padrões, (v) Prevenção de erro, (vi) Reconhecimento em vez de memorização, (vii) Flexibilidade e eficiência de uso, (viii) Estética e *design* minimalista, (ix) Ajudar os usuários a reconhecer, diagnosticar e recuperar de erros e (x) Ajuda e documentação.

As heurísticas foram analisadas considerando graus de severidade da seguinte forma: Grau de Severidade 1 (GS1) - Não concordo que seja um problema de usabilidade; Grau de Severidade 2 (GS2) - Cosmético: Consertar apenas se sobrar tempo; Grau de Severidade 3 (GS3) - Menor: Deve receber baixa prioridade de correção; Grau de Severidade 4 (GS4) - Maior: Deve receber alta prioridade de correção; e Grau de Severidade 5 (GS5) - Catastrófico: Correção de imediato. A Figura 4 apresenta a estrutura da coleta dos dados da avaliação das heurísticas de Nielsen.

	GS1. Não concordo que seja um problema de usabilidade.	GS2. Cosmético: Consertar apenas se sobrar tempo.	GS3. Menor: Deve receber baixa prioridade de correção.	GS4. Maior: Deve receber alta prioridade de correção.	GS5. Catastrófico: Correção de imediato.
H1. Visibilidade do estado do sistema.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
H2. Compatibilidade entre o sistema e o mundo real.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
H3. Controle e liberdade do usuário.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
H4. Consistência e padrões.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
H5. Prevenção de erro.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
H6. Reconhecimento em vez de memorização.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
H7. Flexibilidade e eficiência de uso.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
H8. Estética e “design” minimalista	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
H9. Ajudar os usuários a reconhecer, diagnosticar e recuperar de erros.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
H10. Ajuda e documentação.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 4: Formulário de inspeção com as heurísticas com os graus de severidade.

Os especialistas realizaram as avaliações e foram contabilizados as heurísticas e os graus de severidade com maior frequência, considerando todas as tarefas em ambas as plataformas.

3.2.3 Aplicação do Questionário SUS

Segundo Brooke (1996), o *System Usability Scale* (SUS) pode ser utilizado para medir a usabilidade de *websites*. O SUS consiste em dez questões com cinco opções de escolha e uma escala

⁷ A avaliação heurística tem como base em um conjunto de diretrizes, que descrevem características desejáveis da interação chamada por Nielsen de heurísticas. Nielsen determinou um conjunto inicial de heurísticas, após inspecionar mais de 240 problemas de usabilidade.

que varia entre 1 (Discordo totalmente) e 5 (Concordo totalmente). Neste trabalho, esta análise ocorreu, após o final da execução das sete tarefas de cada plataforma pelos especialistas e visava buscar uma visão geral de experiência deles. A contabilização dos resultados seguiu o método indicado em Brooke (1996):

- Dentre as dez perguntas, as ímpares são questões com aspectos positivos, então quanto maior for a quantidade de respostas nos índices de escala 4 ou 5, melhor será o *Score* do SUS.
- Já as questões pares são consideradas questões com aspectos negativos, portanto, neste caso, as respostas deveriam ficar entre 1 ou 2.
- Para chegar no SUS *Score* foi realizado o seguinte cálculo:

- X = soma os pontos das questões ímpares e subtrai 5;
- Y = soma dos pontos das questões pares e subtrai 25;
- SUS *Score* = soma X com Y , em seguida, multiplica por 2,5.

A função do SUS é demonstrar o desempenho de usabilidade nos aspectos de eficácia, eficiência e facilidade de uso geral. Dessa forma, Sauro (2011) apresentou uma escala para validar os resultados do SUS *Score*, com o objetivo de ter uma melhor interpretação dos pontos e uma melhor significância sobre a qualidade da usabilidade do sistema estudado. A Tabela 1 apresenta essa compilação.

Tabela 1: Categorias de avaliação do SUS *Score*.

SUS <i>Score</i>	Avaliação	Classificação
Valores acima de 80.3	A	Excelente
Entre 68.1 e 80.3	B	Bom
<i>Scores</i> iguais a 68	C	Ok
Entre 51 e 68	D	Ruim
Menor que 51	F	Muito ruim

3.2.4 Questionário de Feedback dos Especialistas

Nesta última etapa, utilizando *Word Cloud* (Heimerl, Lohmann, Lange, & Ertl, 2014), realizou-se a análise textual dos comentários inseridos pelos especialistas em um formulário de *feedback*. As perguntas deste formulário tinham o objetivo de extrair dos especialistas os problemas e as sugestões gerais de melhoria de cada plataforma. Elas foram elaboradas e estruturadas da seguinte maneira: (i) “Fique à vontade para descrever os problemas de usabilidade relacionado à sua experiência com a plataforma de modo geral” e (ii) “Quais seriam suas sugestões de melhoria, de modo geral, para sanar ou minimizar os problemas de usabilidade encontrados na plataforma X”. É importante notar que as respostas foram solicitadas para os especialistas, após a realização das sete tarefas, para uma visão mais ampla de cada plataforma.

3.2.5 Validação

Após a realização da inspeção em cada LMS, os dados coletados nos formulários foram analisados qualitativamente para comparar os sistemas deste estudo (veja Figura 3-E). As análises e resultados são apresentados a seguir.

4 Resultados e Discussões

Os resultados da avaliação e da comparação das plataformas para descobrir problemas de usabilidade estão relatados a seguir. Eles são apresentados em seções de acordo com o método utilizado.

4.1 Resultados dos Testes de Usabilidade

A Figura 5 apresenta as medições de tempo, após a execução de cada tarefa, de acordo com o teste de usabilidade. Estes valores viabilizaram a descrição dos resultados da capacidade de aprendizado, de eficiência e de memorização nos sistemas avaliados.

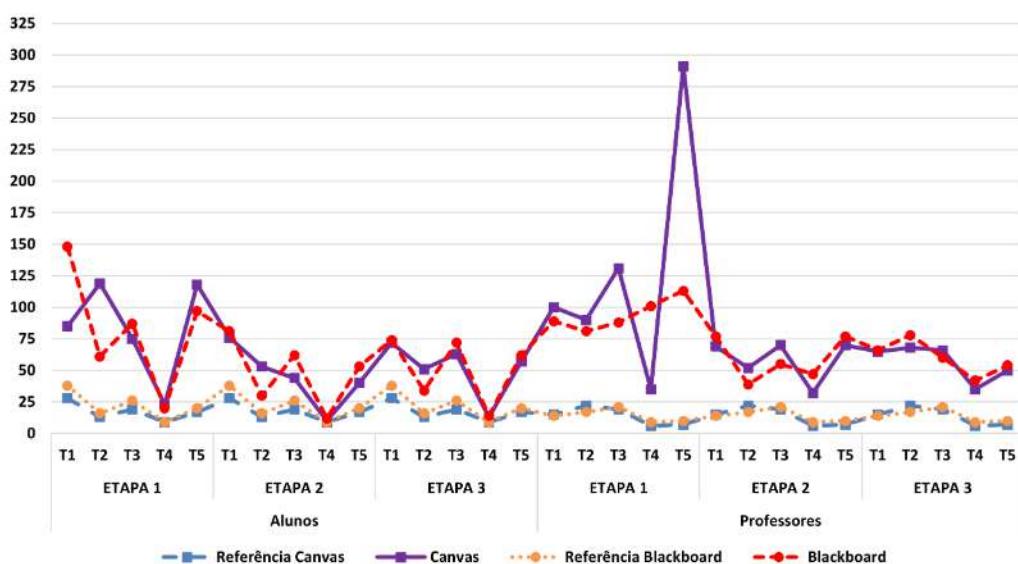


Figura 5: Tempo por tarefa e etapa.

Com relação à capacidade de aprendizado dos alunos, nota-se que as maiores médias de tempo na primeira etapa em comparação com a referência (Seção 3.1.2) foram obtidas nas tarefas 1, 2, 3 e 5 para ambos os sistemas LMS. Em todas essas tarefas, o tempo despendido pelos usuários foi ao menos 2,5 vezes maior que o tempo de referência, sugerindo dificuldade de uso no primeiro acesso aos sistemas. É importante destacar que as tarefas 2 e 5 no Canvas, bem como a tarefa 1 no Blackboard, foram as que alcançaram os maiores tempos de realização. Esses resultados refletem os problemas associados à falta de estruturação das informações, palavras com duplo sentido, e ícones e botões sem a devida padronização, que apenas dificultaram a execução das tarefas.

Já em relação à capacidade de aprendizado do sistema para os professores, destaca-se que as tarefas 1, 3 e 5 no Canvas foram executadas com um tempo ao menos seis vezes maior que a medição de referência. No Blackboard, as tarefas 1, 4 e 5 apresentaram os maiores valores observados, todos ao menos seis vezes maior que a referência.

Os resultados de tempo da segunda etapa indicaram que as tarefas 1 e 2 no Canvas e a tarefa 5 no Blackboard apresentaram problemas de eficiência nos testes de usabilidade com os alunos. Novamente, o tempo despendido por este grupo de usuários foi ao menos 2,5 vezes maior que o tempo de referência nas referidas tarefas. Embora as médias de tempo nesta etapa sejam menores

que os valores observados na primeira etapa, conforme esperado, os resultados apontam que os problemas observados na etapa inicial ainda persistem, mesmo com a repetição das tarefas e as instruções concedidas aos alunos.

Os resultados dos professores em termos de eficiência indicaram que, no Canvas, as tarefas 1, 4 e 5 foram em torno de 60% menos eficazes em relação as demais. Já no sistema Blackboard, as três piores tarefas foram 5, 1 e 3, nesta ordem, apresentando um valor de aproximadamente 90% de diferença para as outras. No geral, o Blackboard apresenta um resultado de eficiência em torno de 8% maior que o Canvas.

Ao analisar as médias de tempo dos alunos na etapa 3, observa-se que nas tarefas 3 e 5 do Canvas ocorreu um aumento de aproximadamente 43% no tempo médio em relação aos valores da etapa 2. No Blackboard, as mesmas tarefas geraram um aumento de aproximadamente 16% no tempo de conclusão observado entre as etapas. Desse modo, os alunos demonstraram dificuldade em memorizar aspectos de uso do Canvas e do Blackboard, após 7 dias sem fazer o uso destes sistemas.

Com relação ao critério de memorização aplicado aos professores, as tarefas 2 e 4 foram aquelas que mais evidenciaram a dificuldade de memorização no Canvas, apresentando uma média de 19,5% de aumento no tempo em relação à etapa 2. No Blackboard, as tarefas 2 e 3 resultaram em maior dificuldade de memorização, com uma média de 55% de acréscimo no tempo de execução quando comparadas com a etapa 2.

A Figura 6 apresenta os resultados da métrica Índice de Sucesso da Tarefa (IST), utilizada para avaliar a efetividade no uso do Canvas e do Blackboard. Com os resultados apresentados, nota-se diferenças sutis e alternadas entre os sistemas, contudo existem grandes diferenças de tempo com a referência criada inicialmente no estudo.

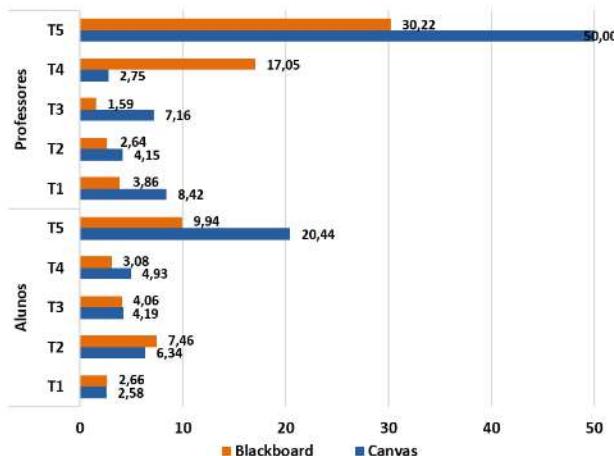


Figura 6: Índice de sucesso na tarefa.

O atributo de efetividade novamente confirmou que ambos os sistemas avaliados apresentaram problemas de usabilidade em diversas tarefas propostas. No Canvas, com exceção da tarefa 1, os alunos apresentaram baixa efetividade em todas as tarefas, resultando em índices de erro no mínimo quatro vezes maior que o desejado. Estes resultados também foram observados no Blackboard, por meio das tarefas 2, 3 e 5, o que significa que os alunos tiveram que realizar inúmeros cliques e trocas de páginas adicionais no navegador para concluir a tarefa proposta.

Já em relação às tarefas realizadas pelos professores, é notório que o Canvas indica ter mais problemas do que o Blackboard, pois apresenta resultados piores nas quatro das cinco tarefas realizadas. Embora a tarefa 5 apresente o pior desempenho em ambos os sistemas, a média de IST no Canvas representa um percentual de diferença para o Blackboard de aproximadamente 350%, já que o valor real de IST obtido pelo Canvas para a tarefa 5 foi de 138 (no gráfico foi apresentado como 50 para facilitar a visualização do comportamento das demais tarefas). Esta diferença foi apontada pelos especialistas principalmente pelos problemas atribuídos à estrutura das páginas (*layout*) de conteúdo que estavam sem organização e com ícones que não correspondiam com a finalidade da sua ação.

O gráfico de radar da Figura 7 ilustra a média das respostas inseridas nos questionários de satisfação, após a conclusão de cada tarefa do teste de usabilidade.

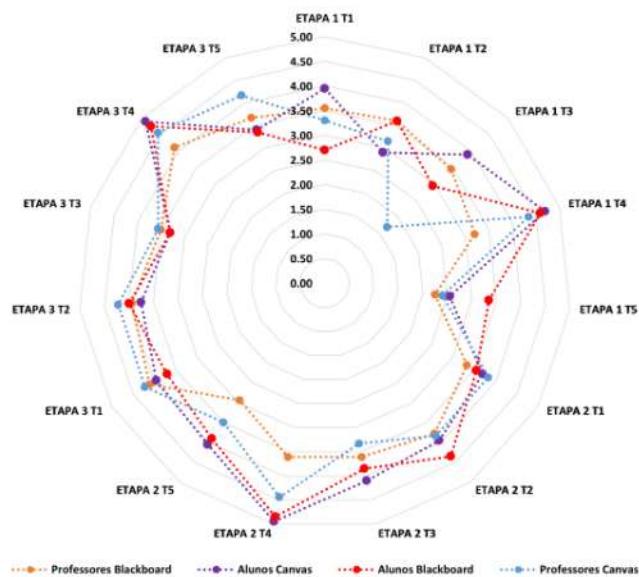


Figura 7: QPC por tarefa e etapa.

A Figura 8 exibe os resultados do questionário de satisfação pós-etapa de acordo com o grupo avaliado no teste de usabilidade.

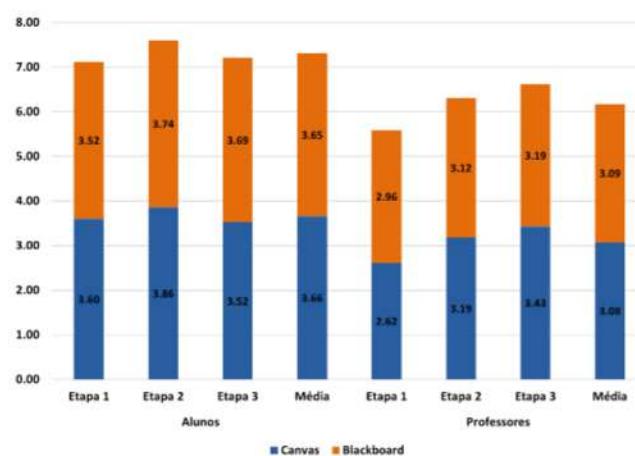


Figura 8: QUSPE por etapa.

Como pode ser observado nas Figuras 7 e 8, inicialmente, os alunos demonstraram-se menos satisfeitos após as tarefas 2 e 5 no Canvas e após as tarefas 1 e 3 no Blackboard. Nas referidas tarefas, os achados foram menores que 3 pontos. Entretanto, logo na etapa posterior, nota-se que os alunos elevaram o grau de satisfação em ambos os sistemas, conforme esperado. Ainda assim, na última etapa, novamente os alunos relataram valores moderados de satisfação para ambos os sistemas, de acordo com os resultados das tarefas 3 e 5.

Finalmente, segundo os resultados dos questionários de satisfação pós-etapa, os sistemas Canvas e Blackboard são considerados semelhantes no atributo de satisfação para o grupo de alunos, uma vez que as médias QPC são equivalentes e as mesmas variações nas opiniões desses usuários são observadas entre as etapas avaliadas.

Para os professores, no sistema Canvas, a tarefa 3 apresentou o maior grau de insatisfação representando em torno de 60% a mais que a média das demais tarefas, porém, ao decorrer das etapas 2 e 3 o valor da tarefa teve uma melhora de 70%, mas ainda abaixo das demais tarefas. Já no sistema Blackboard, o pior grau de satisfação foi atribuído a tarefa 5 na etapa 1, representando uma diferença de 30% em relação as demais tarefas, porém, com a aplicação das etapas 2 e 3, o grau de satisfação se igualou com as outras tarefas.

De modo geral, nas plataformas Canvas e Blackboard, as tarefas 2, 3 e 5 dos alunos estavam relacionadas com mais da metade dos problemas em termos de capacidade de aprender, memorização, eficiência e efetividade.

4.2 Resultados da Inspeção com Especialistas

Para entender os problemas de usabilidade de forma aprofundada nos sistemas, fez-se uma análise com especialistas, que realizaram as mesmas tarefas de alunos e professores para encontrar os problemas e destacar as possíveis soluções, com o intuito de melhorar a usabilidade das plataformas. Os resultados da etapa de inspeção com os especialistas foram organizados de acordo com as subseções a seguir.

4.2.1 Resultados das Heurísticas

Conforme apresentado na Figura 9, especialistas confirmaram a existência de problemas de usabilidade no Blackboard e no Canvas.

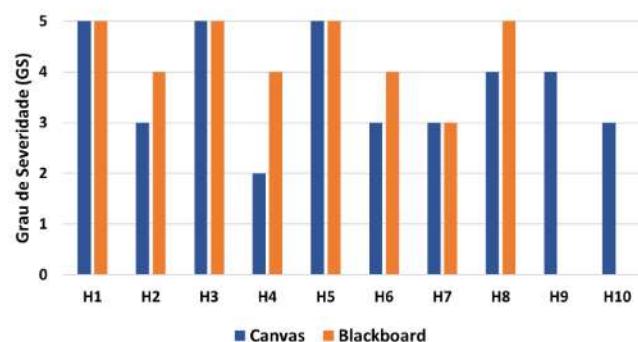


Figura 9: Comparação das plataformas avaliadas em relação às heurísticas.

Ao realizar uma análise minuciosa nas heurísticas em comparação com as plataformas, observa-se que:

- *Visibilidade do estado do sistema (H1)*. Nas duas plataformas, o grau de severidade para essa heurística foi catastrófico (GS5), indicando que os especialistas encontraram problemas graves nesses sistemas. Por exemplo, ambos LMS tinham botões cinzas (com aparência de desabilitados), no entanto, eram funcionais. Essa falta de visibilidade prejudicou muito a realização das tarefas.
- *Compatibilidade entre o sistema e o mundo real (H2)*. Na plataforma Blackboard, essa heurística teve o grau de severidade maior (GS4), enquanto no Canvas, menor (GS3). Os principais problemas relatados foram que os nomes de alguns botões, funcionalidades, ícones não estavam associados ao mundo real e isso confundiu os especialistas, no ato de realizar as tarefas.
- *Controle e liberdade do usuário (H3)*. Nas duas plataformas, o grau de severidade para essa heurística foi catastrófico (GS5) e os principais problemas foram associados à não ter a sensação de controle sob as ações do sistema no ato da realização das tarefas e a dificuldade de cancelar uma ação indevida no sistema.
- *Consistência e padrões (H4)*. Na plataforma Canvas, essa heurística teve o grau de severidade cosmético (GS2) e no Blackboard teve o maior (GS4). Os problemas relatados pelos especialistas foram a falta de padrão que existia entre as janelas na realização das tarefas, cores não relacionadas com as funcionalidades de botões e ainda, sua falta de contrastes. Além disso, a estrutura dos menus não tinha consistência ou padrão.
- *Prevenção de erro (H5)*. Nas duas plataformas, o grau de severidade para essa heurística foi catastrófico (GS5) e os problemas relatados pelos especialistas foram a falta de mensagens de avisos/confirmações em algumas tarefas e filtros nas janelas para impedir a importação de arquivos em formato errado ou não reconhecido pelo sistema.
- *Reconhecimento em vez de memorização (H6)*. Na plataforma Canvas, o grau de severidade para essa heurística foi menor (GS3) e para o Blackboard, maior (GS4). Em ambos, houve muitos problemas de memorização e não remeteram ao reconhecimento. Alguns especialistas relataram que algumas tarefas simples não foram realizadas rapidamente, pois no percurso para atingir o objetivo da atividade houve rótulos em botões ou funcionalidades que não ajudaram no reconhecimento da tarefa.
- *Flexibilidade e eficiência de uso (H7)*. Em ambas as plataformas, o grau de severidade para essa heurística foi menor (GS3) e os principais pontos abordados foram os problemas de flexibilidade, como, por exemplo, na tarefa de enviar uma mensagem, alguns especialistas relataram que não foram utilizados métodos e uma estrutura que sanaria as necessidades de diferentes perfis de usuários.
- *Estética e design minimalista (H8)*. Na plataforma Blackboard, o grau de severidade para essa heurística foi catastrófico (GS5) e no Canvas foi de severidade menor (GS4). Os especialistas relataram que os problemas foram a falta de organização e estruturação das informações nos sistemas e ainda, um excesso de funções não essenciais gerando poluição visual.
- *Ajudar os usuários a reconhecer, diagnosticar e recuperar de erros (H9)*. Nesta heurística, a única plataforma relacionada foi o Canvas, o qual obteve um grau de severidade maior (GS4). Os especialistas constataram que o Canvas, em algumas tarefas, não apresentava solução para resolver ou reconhecer a gravidade do problema encontrado no ato da realização das atividades.
- *Auxiliar e documentação (H10)*. A plataforma Canvas foi a única apontada nessa heurística por alguns especialistas e o seu grau de severidade foi menor (GS3). Os especialistas relataram que não

existe no Canvas um botão ou funcionalidade de ajuda para tirar dúvida. Segundo eles, algumas tarefas sofreram prejuízo para atingir seus objetivos.

De modo geral, pode-se levantar alguns pontos importantes sobre as heurísticas violadas em ambas as plataformas. O primeiro ponto é que o Canvas obteve menos heurísticas afetadas com nível de severidade alto em comparação com o Blackboard. O segundo ponto é que o Blackboard tem menos heurísticas afetadas em qualquer nível de grau de severidade. Já o terceiro ponto é que, em ambas as plataformas, as heurísticas H1, H3 e H5 foram indicadas com o grau de severidade catastrófico (GS5) solicitando uma correção imediata. Com isso, pode-se concluir que Canvas e Blackboard têm seus pontos positivos e negativos, mas que alguns ajustes precisam ser feitos para que esses sistemas possam resolver problemas de usabilidade, principalmente de visibilidade de status, associação de metáforas com mundo real e uso de filtros e mensagens eficientes para prevenção de erros.

4.2.2 Resultados do Questionário SUS

A Figura 10 ilustra a comparação entre as plataformas Canvas e Blackboard, após aplicação do questionário SUS com os especialistas. No eixo horizontal são alocados os especialistas (E1 para Especialista 1, E2 para Especialista 2, etc.). No eixo vertical, os valores de *Score* SUS de cada especialista separado por cada plataforma são pontuados.

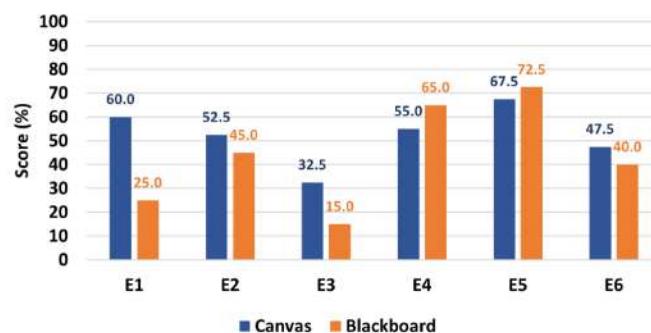
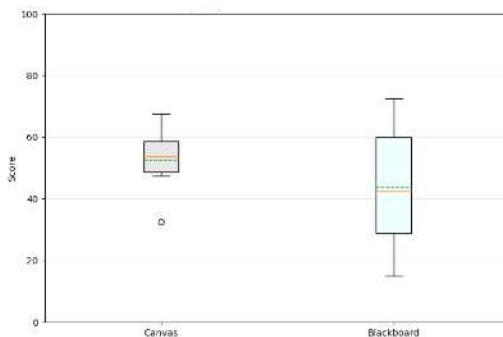


Figura 10: Comparação das plataformas avaliadas em relação ao questionário SUS.

A Figura 11 ilustra a comparação das plataformas utilizando o gráfico de *Boxplot*, para a análise exploratória das variáveis quantitativas. Ele é capaz de fornecer diversas medidas estatísticas como mínimo, máximo, primeiro quartil, segundo quartil (mediana) e terceiro quartil dos *scores* dos especialistas. De modo geral, a pontuação média de satisfação do *Score* do SUS é 68. Analisando os resultados de comparação com as plataformas, pode-se observar que:

- O Blackboard em comparação com o Canvas alcançou um *score* maior do que 68 com o Especialista 5. Contudo, não foi unânime, pois alguns especialistas geraram análises e *scores* diferentes.
- Analisando a média e a mediana, o Canvas obteve maiores resultados que o Blackboard. Entretanto, os resultados de ambos ficaram abaixo do valor 68, que é o parâmetro de satisfação.

O resultado do questionário SUS de comparação das plataformas confirma que ações precisam ser feitas para que as usabilidades de Blackboard e de Canvas sejam aprimoradas, tornando a interação com os sistemas mais consistentes e oferecendo uma melhor experiência aos usuários.

Figura 11: *Boxplot* das plataformas avaliadas com o SUS.Tabela 2: Dados oriundos do *Boxplot*.

LMS	Min	Média	Max	1 quartil (25%)	2 quartil mediana (50%)	3 quartil (75%)
Canvas	32,5	52,5	67,5	48,75	53,75	58,75
BB	15	43,75	72,5	28,75	42,5	60

4.2.3 Resultados do Questionário de Feedback dos Especialistas

De análise textual dos questionários, encontrou-se que as palavras mais importantes relacionadas aos problemas encontrados pelos especialistas no Canvas foram: **ícone, memorização, pena, botão e confusão**. Todas essas palavras, segundo os especialistas, estão relacionadas com todas as tarefas. Foram abordados problemas como botões não representativos com o mundo real (como o ícone de “criar um e-mail” que era uma “pena”), a interação com a ferramenta e suas funcionalidades. Foi muito questionado que, ao realizar as tarefas, o princípio partia mais por memorização, e isso, em um contexto geral, causou certa confusão por falta de consistência.

De modo geral, os especialistas, em sua grande maioria, propuseram uma mudança de estrutura e de organização do painel lateral, do ambiente de enviar mensagens buscando maiores funcionalidades e nomes mais relacionados as suas funções, por exemplo, botão “pessoa” trocar para “alunos”. Uma outra sugestão foi diminuir as informações ou adicioná-las em *cards* para cada tópico principal das páginas, tornando a aprendizagem e a visualização do sistema menos poluída para menores sobrecargas cognitivas. Outro ponto relatado foi melhorar a acessibilidade do sistema e buscar novas formas de estruturá-lo com mais métodos de reconhecimento para atingir todos os públicos-alvo da plataforma.

Com relação aos problemas do Blackboard, as palavras com maiores frequências foram: **botão, notas e estrutura**. Nota-se que a relação entre elas está nas tarefas 2, 3, 6 e 7. Os especialistas relataram que não foi possível reconhecer que alguns botões tivessem relação com as atividades, pois, faltou contraste e isso atrapalhou na realização das tarefas. Além disso, relataram a dificuldade de reconhecer algumas funcionalidades no sistema, e ainda, as grandes estruturas de algumas páginas (*layout*) de conteúdo, deixando sem consistência e padrão a interação com a plataforma.

As soluções para a plataforma Blackboard, conforme os especialistas, de modo geral, indicaram mudanças que poderiam tornar o sistema flexível e, consequentemente, de melhor usabilidade

para os usuários. As propostas de maior destaque foram buscar um *design* mais minimalista, reduzir informação desnecessária dos *layouts*, rever conceitos e nomes de botões com a utilização de contrastes maiores para representar que têm funcionalidade neles e ainda, mudar a estrutura do menu lateral. Outro ponto foi adicionar conceitos de *breadcrumbs*, que representam o caminho percorrido pelo usuário em um sistema para facilitar a navegação e tornar o uso mais prático na plataforma.

Em ambas as plataformas, os problemas recorrentes são a falta de organização do conteúdo, informação desnecessária que pode causar confusão na interação e na visualização dos sistemas e problemas nos “nomes” e figuras dos botões com falta de representatividade. As soluções desses problemas apontados pelos especialistas são sugeridas como recomendações, podendo gerar maior interação entre o usuário e os sistemas e consequente maior usabilidade, satisfação e eficiência.

5 Considerações Finais

Este artigo teve como finalidade explorar, avaliar, comparar e propor soluções que envolvam a usabilidade de duas plataformas de ensino à distância do tipo LMS. Diversos trabalhos foram encontrados na literatura em termos de plataformas LMS e avaliação de usabilidade. Contudo, nenhum dos estudos citados abordou a avaliação de usabilidade do Canvas e Blackboard utilizando conjuntamente usuários (alunos e professores) e especialistas da área de IHC. Além disso, diferentemente de Gumasing et al. (2022), que avaliou a usabilidade destes sistemas considerando apenas a satisfação de alunos com o método SUS, o presente estudo aplicou uma abordagem de avaliação aprofundada, considerando testes de usabilidade com usuários, questionários de satisfação, bem como, inspeção com especialistas para calcular a pontuação do método SUS e apontar os problemas e soluções mediante as heurísticas de Nielsen.

A administração da universidade avalia que o impacto da migração dos sistemas LMS foi positivo considerando custos mais baixos, principalmente, devido aos termos da nova contratação, à estabilidade, à ilimitação de horários para uso de serviços e o suporte de alta demanda de requisições.

Sumarizando os resultados encontrados com os experimentos do artigo, nota-se que ambos os sistemas, na percepção dos alunos e professores, tiveram resultados similares em relação aos atributos de usabilidade avaliados, tanto em termos quantitativos como qualitativos. Estes resultados estão de acordo com os achados de Gumasing et al. (2022) e dificultam o apontamento do sistema de melhor usabilidade, mas permitem compreender que alunos e professores não tiveram ganho significativo de usabilidade após a implantação do novo sistema. No entanto, é importante destacar que o trabalho proporciona uma visão geral das dificuldades encontradas na realização das tarefas especificadas, que possuem características semelhantes em ambas as plataformas. Desta forma, apresentou-se um mapeamento dos problemas.

Dado que os fluxos realizados em ambos os sistemas LMS avaliados resultaram em diversos problemas de usabilidade, as devidas correções devem ser planejadas e implementadas como estratégia de intervenção, principalmente no sistema Blackboard, atualmente utilizado pela instituição. De acordo com um estudo realizado por Oliveira et al. (2020), a maioria dos alunos discorda quanto às questões referentes a “Facilidade de Uso Percebida” no Blackboard, resultando

em desmotivação para utilizar o AVA. Finalmente, apesar de existirem alguns casos de sucesso de sistemas LMS no Brasil, como o Moodle da Universidade Estadual do Maranhão (Brasil, 2022), muitos usuários acabam utilizando o AVA por obrigação e não voluntariamente (Oliveira et al., 2020).

As principais contribuições do presente artigo são:

- o apontamento dos problemas mais relevantes encontrados em ambas as plataformas com propostas de possíveis soluções para o aprimoramento das plataformas em relação à usabilidade. Salienta-se que as indicações providas pelos especialistas vão além do fluxo das tarefas e podem ser aplicadas nos LMS em termos de serviços da plataforma e de itens configuráveis da interface de usuário. Em ambos os casos, os resultados e as sugestões apontam para possíveis pontos de atenção no *design* dos LMS e para oportunidades de pesquisar o impacto dos problemas ou das melhorias sugeridas. Em suma, aspectos como consistência, padrão, organização e *design* minimalista do *layout* são fundamentais para um LMS, ao passo que devem evitar sobrecarga de informações, botões e nomes não-representativos com o mundo real, funcionalidades não explícitas ao usuário, e ainda, cores e tamanhos inadequados.
- a apresentação detalhada de uma metodologia que pode ser reutilizada para avaliações de usabilidade, principalmente de sistemas do tipo LMS que estão em expansão devido aos contextos atuais da educação em constante mudança. A mesma metodologia pode ser usada para avaliar os sistemas de maneira unitária ou comparativa. A adaptação de parâmetros e participantes para este estudo foi muito satisfatória.
- apesar de ser um estudo de caso, os resultados do trabalho podem contribuir para o entendimento dos sistemas avaliados, provendo indagações sobre a extensão da aplicabilidade dos resultados e abrindo oportunidade de pensar sobre melhorias de eficiência, de eficácia e de continuidade de uso desses sistemas.

Existem outros trabalhos futuros para continuação da investigação que são:

- explorar outras diferentes tarefas em ambos os LMS, uma vez que as tarefas avaliadas foram representativas, mas simples em termos da sequência de passos para suas execuções. Essas tarefas devem ser mais robustas, principalmente, para os professores que devem realizar atividades de incluir conteúdos e de avaliação;
- criar avaliações independentes para as funcionalidades providas pelos LMS e para os itens configuráveis de *design* das interfaces dos usuários; e
- realizar novas avaliações com a alteração dos parâmetros de aplicação das etapas para observar diretamente os critérios de usabilidade (memorização e capacidade de aprender).

Agradecimentos

Os autores agradecem à Universidade Virtual do Estado de São Paulo (UNIVESP) pelo apoio financeiro concedido durante o desenvolvimento deste trabalho. Agradecemos também à Universidade de São Paulo (USP), à Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP) e à Universidade Estadual de Campinas (UNICAMP) pelos convênios realizados com a UNIVESP que possibilitaram o desenvolvimento deste estudo.

Referências

- Ahmad, F., Beyene, W., & Giannoumis, G. A. (2018). Comparative Evaluation of Accessibility and Learnability of Learning Management Systems: Case of Fronter and Canvas. In C. Stephanidis (Ed.), *Hci international 2018 – posters' extended abstracts* (pp. 3–9). Cham: Springer International Publishing. doi: [10.1007/978-3-319-92279-9_1](https://doi.org/10.1007/978-3-319-92279-9_1) [GS Search]
- Alhadreti, O. (2020). A comparative usability study of blackboard and desire2learn: Students' perspective. In P. Zaphiris & A. Ioannou (Eds.), *Learning and collaboration technologies. designing, developing and deploying learning experiences* (pp. 3–19). Cham: Springer International Publishing. doi: [10.1007/978-3-030-50513-4_1](https://doi.org/10.1007/978-3-030-50513-4_1) [GS Search]
- Alturki, U., Aldraiweesh, A., & Kinshuk. (2016). Evaluating the usability and accessibility of lms “blackboard” at king saud university. *Contemporary Issues in Education Research (CIER)*, 9(1), 33-44. doi: [10.19030/cier.v9i1.9548](https://doi.org/10.19030/cier.v9i1.9548) [GS Search]
- Andrews-Todd, J., Bond, G., Speller, L., Stevenson, R., & Tillman, B. (2013). Usability of the blackboard course management system: Student and instructor dynamics in a dynamic system. *Cognitive Technology Journal*, 18(2), 18-32. [GS Search]
- Associação Brasileira de Normas Técnicas (2002). *NBR 9241-11. Requisitos ergonômicos para trabalho de escritório com computadores: parte 11 - orientação sobre usabilidade*. Rio de Janeiro: ABNT.
- Brasil (2022). *ReUni Digital. Panorama da EaD no Brasil* (Vol. 2). Disponível em [\[Link\]](#).
- Brooke, J. (1996). Sus: A ‘quick and dirty’ usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (1st ed., pp. 1–6). Boca Raton, FL: CRC Press. doi: [10.1201/9781498710411-35](https://doi.org/10.1201/9781498710411-35) [GS Search]
- Chen, W., Sanderson, N. C., Nichshyk, A., Bong, W. K., & Kessel, S. (2021). Usability of learning management systems for instructors – the case of canvas. In P. Zaphiris & A. Ioannou (Eds.), *Learning and collaboration technologies: New challenges and learning experiences* (Vol. 12784, pp. 210–223). Cham: Springer International Publishing. doi: [10.1007/978-3-030-77889-7_14](https://doi.org/10.1007/978-3-030-77889-7_14) [GS Search]
- Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments and Computers*, 35(3), 379-383. doi: [10.3758/bf03195514](https://doi.org/10.3758/bf03195514) [GS Search]
- Grossi, M. G. R., de Souza Elias, M. C. A., Chamon, C. M., & Leal, D. C. C. C. (2018). The educational potentialities of the virtual learning environments moodle and canvas: A comparative study. *International Journal of Information and Education Technology*, 8(7), 514-519. doi: [10.18178/ijiet.2018.8.7.1091](https://doi.org/10.18178/ijiet.2018.8.7.1091) [GS Search]
- Gumasing, M. J. J., Vasquez, A. B., Doctora, A. L. S., & Perez, W. D. D. (2022). Usability evaluation of online learning management system: Comparison between blackboard and canvas. In *2022 the 9th international conference on industrial engineering and applications (europe)* (p. 25–31). New York, NY, USA: Association for Computing Machinery. doi: [10.1145/3523132.3523137](https://doi.org/10.1145/3523132.3523137) [GS Search]
- Gunasekera, A. I., Bao, Y., & Kibelloh, M. (2019). The role of usability on e-learning user interactions and satisfaction: a literature review. *Journal of Systems and Information Technology*, 21(3), 368-394. doi: [10.1108/JSIT-02-2019-0024](https://doi.org/10.1108/JSIT-02-2019-0024) [GS Search]
- Heimerl, F., Lohmann, S., Lange, S., & Ertl, T. (2014). Word cloud explorer: Text analytics based on word clouds. In *2014 47th hawaii international conference on system sciences*

- (p. 1833-1842). Waikoloa, HI, USA: IEEE. doi: [10.1109/HICSS.2014.231](https://doi.org/10.1109/HICSS.2014.231) [GS Search]
- Lewis, J. R. (1995). Ibm computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human–Computer Interaction*, 7(1), 57-78. doi: [10.1080/10447319509526110](https://doi.org/10.1080/10447319509526110) [GS Search]
- Nielsen, J. (1994). Heuristic evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability inspection methods* (1st ed., p. 25–62). New York, NY, USA: John Wiley & Sons, Inc.
- Nielsen, J. (2012, jun). *How many test users in a usability study?* Disponível em [\[Link\]](#).
- Nielsen, J., & Loranger, H. (2006). *Prioritizing web usability*. Thousand Oaks, CA, USA: New Riders Publishing.
- Oliveira, N. J. S., Procaci, T. B., & Siqueira, S. W. M. (2020). Captura da Aceitação do Blackboard e do Tipo de Motivação de Alunos de Cursos Presenciais de Ciências Exatas em uma Universidade Privada. *Revista Brasileira de Informática na Educação*, 28, 229-259. doi: [10.5753/RBIE.2020.28.0.229](https://doi.org/10.5753/RBIE.2020.28.0.229) [GS Search]
- Phongphaew, N., & Jiamsanguanwong, A. (2018). Usability evaluation on learning management system. In T. Ahram & C. Falcão (Eds.), *Advances in usability and user experience* (pp. 39–48). Cham: Springer International Publishing. doi: [10.1007/978-3-319-60492-3_4](https://doi.org/10.1007/978-3-319-60492-3_4) [GS Search]
- Rocha, H. V., & Baranauskas, M. C. (2003). Avaliação de interfaces. In *Design e avaliação de interfaces humano-computador* (p. 200-208). Campinas: UNICAMP.
- Sauro, J. (2011). *A practical guide to the system usability scale: Background, benchmarks & best practices*.
- Sobaih, A. E. E., Hasanein, A. M., & Abu Elnasr, A. E. (2020). Responses to covid-19 in higher education: Social media usage for sustaining formal academic communication in developing countries. *Sustainability*, 12(16), 6520. doi: [10.3390/su12166520](https://doi.org/10.3390/su12166520) [GS Search]