UNIVERSIDADE DE BRASÍLIA
FACULDADE UNB GAMA
Disciplina: Interação Humano-Computador
Professor: André Barros de Sales

VICTOR ANDREOZZI DE LA ROCQUE COUTO
Matrícula: 221029285

ENTREGA 7: RELATO DOS RESULTADOS DA AVALIAÇÃO DO PROTÓTIPO DE ALTA FIDELIDADE

Parte individual: Avaliação de Disciplinas

Projeto No Fluxo UnB · Grupo 02 · 2026.2

Vídeo da apresentação: https://www.youtube.com/watch?v=vO76Ck3ILlQ

BRASÍLIA
2026

SUMÁRIO

1 OBJETIVOS E ESCOPO DA AVALIAÇÃO

2 MÉTODO: COMO A AVALIAÇÃO FOI REALIZADA

3 AVALIADOR E PARTICIPANTES

4 TAREFAS EXECUTADAS PELOS PARTICIPANTES

5 MEDIÇÕES REALIZADAS

6 INTERPRETAÇÃO E ANÁLISE DOS DADOS

6.1 Análise de cada participante (intraparticipante)

6.2 Consolidação entre participantes (interparticipante)

6.3 Respostas às perguntas da avaliação

6.4 Confiabilidade e validade

7 PROBLEMAS ENCONTRADOS

8 PLANEJAMENTO PARA O REPROJETO

9 CONCLUSÃO

10 CUIDADOS ÉTICOS

11 DECLARAÇÃO SOBRE O USO DE IA GENERATIVA

12 REFERÊNCIAS

Este documento relata os resultados do teste de usabilidade do Protótipo de Alta Fidelidade da Avaliação de Disciplinas, conforme o planejamento do relato da Entrega 6. A estrutura segue os itens que Barbosa e Silva (2010) indicam para o relato de um teste de usabilidade (seção 10.2.1) e de qualquer avaliação de IHC (seção 9.7.5): objetivos e escopo, método, avaliador e participantes, tarefas, tabelas e gráficos das medições, interpretação e análise, lista de problemas e planejamento do reprojeto.

1 OBJETIVOS E ESCOPO DA AVALIAÇÃO

Em um teste de usabilidade, os objetivos determinam quais critérios de usabilidade são medidos, e cada critério é explorado por perguntas associadas a um dado mensurável (Barbosa e Silva, 2010, seção 10.2.1). A avaliação teve três objetivos, definidos no planejamento com o framework DECIDE (Barbosa e Silva, 2010, seção 9.8):

- medir a eficácia, a eficiência e a satisfação dos estudantes ao avaliar e consultar uma disciplina no protótipo interativo, comparando os resultados com as metas de usabilidade;
- verificar se as correções feitas a partir do protótipo de papel (PP-01 a PP-06) resolveram os problemas e se os problemas das avaliações anteriores (A-01 a A-03 e B-01 a B-07) não reapareceram;
- identificar problemas de fluxo, navegação e aparência que o protótipo de papel não podia revelar, como rolagem, estados de envio e descarte e consistência visual.

O escopo é o protótipo vertical da Avaliação de Disciplinas, construído em HTML, CSS e JavaScript: busca e resultados (T1 e T2), página da disciplina (T3), formulário de avaliação (T4), revisão (T5) e confirmação (T6), com os estados de envio e descarte (Preece, Rogers e Sharp, 2005, seção 8.2.5). As oito perguntas da avaliação são respondidas na Tabela 9, como pedem Barbosa e Silva (2010, seção 9.7.5).

2 MÉTODO: COMO A AVALIAÇÃO FOI REALIZADA

O método foi o teste de usabilidade (Rubin, 1994; Rubin e Chisnell, 2008, citados por Barbosa e Silva, 2010, seção 10.2.1): usuários-alvo realizaram tarefas no protótipo enquanto eram registrados dados sobre o seu desempenho e as suas opiniões. Cada sessão seguiu a coleta de dados descrita por Barbosa e Silva (2010, seção 9.7.3): apresentação dos objetivos e do consentimento, sessão de observação com a técnica de pensar em voz alta (*think aloud*) e uma entrevista pós-teste curta sobre a experiência. A tela, o áudio e o vídeo foram gravados, e o avaliador reviu as gravações para medir tempos, caminhos e falas.

Barbosa e Silva (2010, seção 9.7.5) observam que os resultados são fortemente influenciados pelo ambiente de avaliação, por isso o relato registra como a avaliação foi de fato realizada. A Tabela 1 compara o planejamento, em sua versão ajustada, com o que foi realizado. A diferença que restou está nas tarefas de P1 a P3.

Tabela 1: Forma da avaliação: planejado e realizado

Item	Planejado (Entrega 6, versão ajustada)	Realizado	Efeito nos resultados
Data e horário	12/11 e 14/11/2026	P1 a P3 em 12/11/2026 (8h53, 9h14 e 9h45); P4 e P5 em 14/11/2026 (19h41 e 19h55), horário de Brasília	Como planejado.
Duração	Cerca de 10 minutos por sessão	Gravações de 3 min 14 s a 6 min 5 s	Como planejado; a entrevista pós-teste foi breve.
Consentimento	TCLE mostrado ao participante e aceite registrado na gravação	Como planejado, com o direito de sair a qualquer momento informado	O consentimento foi registrado, mas sem via assinada.
Gravação	Reunião do Microsoft Teams com tela, áudio e câmera	Como planejado, com o avaliador ao lado do participante	O rosto aparece nos vídeos, que ficam com acesso restrito, fora do repositório público.
Perfil	Curso registrado pelo avaliador	Como planejado (Tabela 2)	Semestre e situação no fluxo não foram coletados.
Fala durante as tarefas	Pensar em voz alta	Como planejado	As falas explicam cada passo, mas aumentam o tempo das tarefas (Barbosa e Silva, 2010, seção 9.7.3).
Tarefas	Três cenários com disciplinas escolhidas pelo participante: avaliar, buscar e avaliar, começar e descartar	P4 e P5 fizeram os três cenários. P1 a P3 fizeram só uma tarefa aberta, avaliar uma disciplina já cursada	A Tarefa 1 foi medida com os cinco, e as Tarefas 2 e 3 só com P4 e P5.
Pós-teste	Entrevista com perguntas abertas	Perguntas sobre o que confundiu, o que poderia melhorar e acessibilidade	A satisfação é analisada pelas falas (Tabela 8).
Avaliador	Um, que revê as gravações depois	Victor conduziu as cinco sessões sozinho e reviu e transcreveu todas as gravações	Como planejado.

Fonte: Victor Andreozzi de La Rocque Couto (2026), a partir das gravações das sessões.

Deixar o participante escolher disciplinas que realmente cursou aproximou o teste do uso real: cada um avaliou experiências suas, e não um cenário inventado. Em contrapartida, com P1 a P3, as tarefas foram apresentadas como uma só instrução aberta; a partir de P4, os três cenários foram apresentados um a um. O pensar em voz alta foi útil para a análise qualitativa, mas, como alertam Barbosa e Silva (2010, seção 9.7.3), o participante pode gastar mais tempo falando e pensando, o que deve ser considerado ao comparar os tempos com as metas.

3 AVALIADOR E PARTICIPANTES

O avaliador foi Victor Andreozzi de La Rocque Couto, que apresentou o teste, observou, conduziu a entrevista pós-teste e depois reviu as gravações. Os participantes são identificados por códigos, na ordem em que as sessões ocorreram, sem nome nem matrícula. Nenhum deles participou do teste do protótipo de papel, como previa o planejamento, e o grupo teve três homens e duas mulheres.

Tabela 2: Perfil dos participantes

Código	Sessão	Gênero	Curso	Disciplinas avaliadas no teste
P1	12/11/2026, 8h53	Masculino	Engenharia de Software	Interação Humano-Computador (FGA0210) e Teoria Política Moderna (IPO0205)
P2	12/11/2026, 9h14	Masculino	Engenharia de Software	Orientação a Objetos (FGA0158)
P3	12/11/2026, 9h45	Feminino	Engenharia de Software	Cálculo 1 (MAT0025)
P4	14/11/2026, 19h41	Masculino	Audiovisual	Teorias da Comunicação (FAC0115) e Cálculo 2 (MAT0026); começou e descartou Introdução à Álgebra Linear (MAT0031)
P5	14/11/2026, 19h55	Feminino	Psicologia	Psicologia do Trabalho (PST0001) e Psicologia Cognitiva 1 (PPB0046); começou e descartou Psicopatologia 1 (PCL0021)

Fonte: Victor Andreozzi de La Rocque Couto (2026).

Todos são estudantes de graduação da UnB que já cursaram as disciplinas que avaliaram. P1, P2 e P3 cursam Engenharia de Software, P4 cursa Audiovisual e P5 cursa Psicologia. A variedade de cursos permitiu testar a busca e a avaliação com disciplinas de áreas distintas (engenharia, comunicação, ciência política e psicologia), e não só com as do curso do avaliador. O número de cinco participantes segue Nielsen (1993), segundo quem cinco usuários revelam cerca de 85% dos problemas de usabilidade (Santa Rosa e Moraes, 2008, p. 156 e 157); Barbosa e Silva (2010, seção 9.7.2) também citam esse número para avaliações com observação de usuários.

4 TAREFAS EXECUTADAS PELOS PARTICIPANTES

Barbosa e Silva (2010, seção 10.2.1) pedem que o relato descreva as tarefas executadas. A Tabela 3 mostra as tarefas planejadas na Tabela 12 do planejamento e como foram aplicadas. O caminho mais eficiente previsto na HTA, usado para contar os desvios, é: tela inicial, Avaliações, encontrar a disciplina, Avaliar, preencher as escalas, Revisar e enviar, Confirmar envio.

Tabela 3: Tarefas planejadas e aplicadas

Tarefa	**Cenário planejado**	**Cenário aplicado**	**Participantes**	**Perguntas investigadas**
1 Avaliar uma disciplina	Avaliar uma disciplina já cursada, escolhendo-a direto na lista, sem busca.	“Avalie uma disciplina que você já cursou.” Para P5, “sem fazer nenhuma busca”.	P1 a P5	1, 2, 3, 4 e 5
2 Buscar e avaliar	Buscar pelo nome uma disciplina do seu curso e avaliá-la.	Avaliar uma segunda disciplina; P5 “buscando por uma disciplina do seu curso” e P4 “sem digitar”, escolhendo na lista.	P4 e P5 (P1 fez uma segunda avaliação por iniciativa própria)	3 e 6
3 Desistir da avaliação	Começar a avaliar outra disciplina, parar no meio e descartar.	“Comece uma avaliação, pare no meio e descarte tudo o que fez.”	P4 e P5	7

Fonte: Victor Andreozzi de La Rocque Couto (2026), a partir das gravações das sessões.

5 MEDIÇÕES REALIZADAS

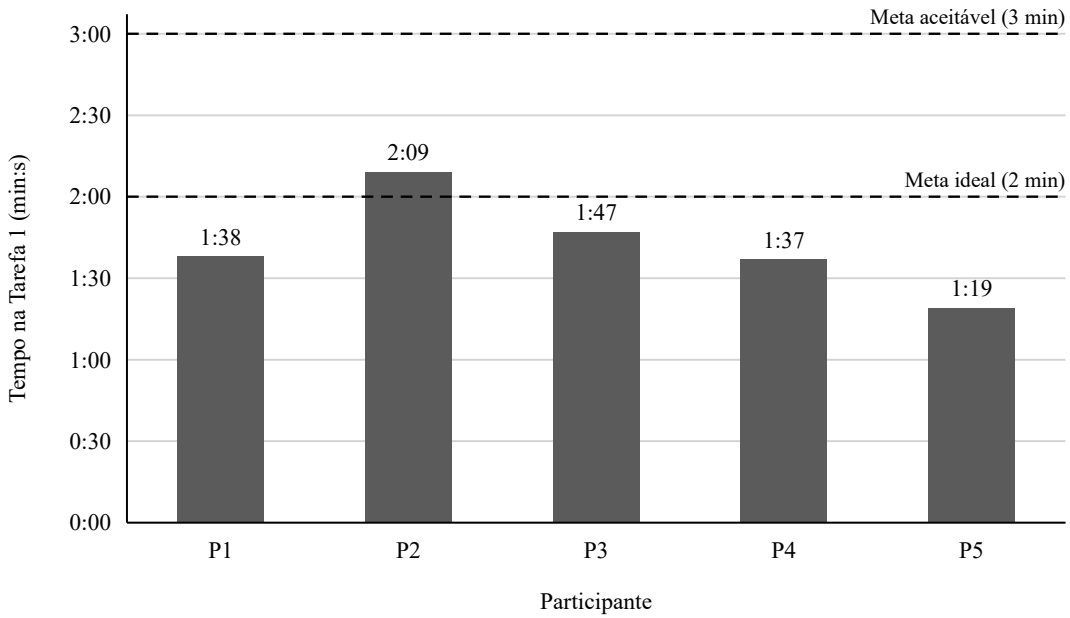
Para cada tarefa de cada participante, foram medidos o grau de sucesso, os erros e desvios, os pedidos de ajuda e o tempo, dados que Barbosa e Silva (2010, seção 10.2.1) citam como típicos do teste de usabilidade. Os tempos foram medidos nas gravações pela mudança de tela registrada na barra de endereço do protótipo, até a tela “Avaliação concluída”. O início foi o aparecimento da tela inicial (P1 a P3) ou o fim das instruções do avaliador, quando a tela inicial já estava aberta durante a explicação (P4 e P5). As medidas são apresentadas por participante, com média e desvio padrão, e comparadas às metas da Tabela 11 do planejamento (meta ideal de até 2 minutos e aceitável de até 3 minutos), como no exemplo de Barbosa e Silva (2010, seção 10.2.1).

Tabela 4: Desempenho de cada participante na Tarefa 1

Participante	Resultado	Tempo total	Como encontrou a disciplina	Desvios do caminho previsto	Pedidos de ajuda	Avaliação registrada
P1	Concluiu sem ajuda	1 min 38 s	Abriu o filtro de professor para conhecer as opções, digitou “int” na busca e abriu “Ver avaliações” antes de avaliar.	0 (a consulta foi uma escolha, e não um erro)	0	Dificuldade 5 e cobrança 5, semestre 2025.2, com professor e comentário.
P2	Concluiu sem ajuda	2 min 9 s	Rolou a lista até Orientação a Objetos, sem usar a busca, e clicou em “Avaliar”.	0	0	Dificuldade 3 e cobrança 2, semestre 2025.1, “Prefiro não informar” o professor, com comentário.
P3	Concluiu sem ajuda	1 min 47 s	Clicou em “Avaliar” em Cálculo 1, primeiro item da lista.	0	0	Dificuldade 4 e cobrança 3, semestre 2026.1, com professor e comentário.
P4	Concluiu sem ajuda	1 min 37 s	Digitou “comuni” na busca, abriu “Ver avaliações” de Teorias da Comunicação e depois “Avaliar esta disciplina”.	0 (a consulta foi uma escolha)	0 (as perguntas feitas foram sobre o teste, antes de começar)	Dificuldade 3 e cobrança 4, semestre 2024.2, com professor e comentário.
P5	Concluiu sem ajuda	1 min 19 s	Digitou “psci” na busca e clicou em “Avaliar” em Psicologia do Trabalho.	0	0	Dificuldade 2 e cobrança 3, semestre 2026.1, “Prefiro não informar” o professor, com comentário.

Fonte: Victor Andreozzi de La Rocque Couto (2026), a partir das gravações das sessões.

Figura 1: Tempo de cada participante na Tarefa 1, comparado às metas



Fonte: Victor Andreozzi de La Rocque Couto (2026). As linhas tracejadas marcam as metas da Tabela 11 do planejamento.

Tabela 5: Eficácia e eficiência por tarefa

Tarefa	Sem ajuda	Com ajuda	Não concluíram	Tempo médio (desvio padrão)	Erros e desvios (total)	Meta atingida?
1 Avaliar uma disciplina	5 de 5	0 de 5	0 de 5	1 min 42 s (± 18 s)	0	Sucesso: ideal. Tempo: média na meta ideal; 4 de 5 na meta ideal e 5 de 5 na aceitável.
2 Buscar e avaliar	2 de 2	0 de 2	0 de 2	43 s (± 16 s)	0	Sucesso: ideal.
3 Desistir da avaliação	2 de 2	0 de 2	0 de 2	29 s (± 16 s)	0 envios acidentais	Ideal para os dois participantes que a realizaram.

Fonte: Victor Andreozzi de La Rocque Couto (2026). Desvio padrão amostral. Os tempos incluem as falas do pensar em voz alta. Tarefas 2 e 3 medidas com P4 e P5 (Tabela 7).

Para entender onde o tempo foi gasto, a Tarefa 1 foi dividida em quatro etapas, também medidas pela mudança de tela (Tabela 6). O preenchimento do formulário foi a etapa mais longa para todos, o que é esperado, porque inclui escolher o semestre e o professor, dar as duas notas e escrever o comentário. A revisão foi curta, com exceção de P2, que leu em voz alta cada item antes de confirmar; P5 confirmou em menos de 4 segundos, sem ler a revisão.

Tabela 6: Tempo por etapa da Tarefa 1

Etapa (telas)	P1	P2	P3	P4	P5	Média
Até a página de Avaliações	7 s	26 s	17 s	1 s	5 s	11 s
Encontrar a disciplina e abrir o formulário (T1 a T3)	31 s	20 s	12 s	38 s	18 s	24 s
Preencher o formulário (T4)	47 s	56 s	70 s	41 s	52 s	53 s
Revisar e confirmar (T5)	13 s	27 s	9 s	17 s	4 s	14 s
Total	**1:38**	**2:09**	**1:47**	**1:37**	**1:19**	**1:42**

Fonte: Victor Andreozzi de La Rocque Couto (2026), a partir das gravações das sessões.

Tabela 7: Desempenho nas Tarefas 2 e 3 e nas avaliações seguintes

Participante	Tarefa	O que fez	Tempo	Resultado
P4	2 Avaliar outra disciplina, sem digitar	Clicou em “Avaliar” em Cálculo 2 na lista, preencheu, revisou e confirmou: “segue o mesmo caminho”.	32 s	Concluiu sem ajuda.
P5	2 Buscar e avaliar	Buscou Psicologia Cognitiva 1, preencheu, revisou e confirmou.	54 s	Concluiu sem ajuda.
P4	3 Desistir	Abriu Introdução à Álgebra Linear, escolheu o semestre, o professor e a dificuldade, clicou em “Descartar” e confirmou no diálogo; leu em voz alta que o rascunho seria perdido e que nada foi enviado.	18 s	Saiu sem enviar, sem ajuda.
P5	3 Desistir	Buscou Psicopatologia 1, deu as notas e hesitou (“E aí eu paro aqui? Eu só saio?”) antes de escolher “Descartar” e confirmar.	40 s	Saiu sem enviar, sem ajuda, com uma breve hesitação.
P1	Segunda avaliação, por iniciativa própria	Clicou em “Avaliar outra disciplina” e avaliou Teoria Política Moderna sem comentário, “que é opcional”.	cerca de 40 s	Concluiu sem ajuda.

Fonte: Victor Andreozzi de La Rocque Couto (2026), a partir das gravações das sessões.

Os três participantes que avaliaram uma segunda disciplina levaram bem menos tempo na segunda vez: P1 passou de 1 min 38 s para cerca de 40 s, P4 de 1 min 37 s para 32 s e P5 de 1 min 19 s para 54 s. A queda sugere que o fluxo é aprendido logo no primeiro uso, o que é coerente com a meta de facilidade de aprendizado da Entrega 3.

Como previsto no planejamento, a satisfação é analisada pelas falas espontâneas e pela entrevista pós-teste, que Barbosa e Silva (2010, seção 9.7.3) descrevem como a oportunidade de coletar a opinião do participante sobre a experiência que acabou de viver e de esclarecer suas intenções e interpretações durante as tarefas (Tabela 8).

Tabela 8: Falas sobre a experiência de uso

Participante	Momento	Fala (resumida)	Leitura
P1	Depois da primeira avaliação	Voltou às avaliações da disciplina e comentou que a sua já aparecia; em seguida, avaliou outra disciplina sem comentário, “que é opcional”.	Conferiu a publicação; entendeu que o comentário é opcional.
P1	Pós-teste	Disse que explorou tudo o que tinha interesse, que achou a funcionalidade bem bacana, mais do que imaginava, e que gostou bastante da ideia.	Satisfação e utilidade percebida altas.
P2	Tela de confirmação	Disse que deu certo e que as informações apareceram, “só não apareceu aqui o comentário”; depois encontrou o comentário em “Ver avaliações”.	Dúvida sobre o envio do comentário (AF-01).
P2	Pós-teste	Achou muito bom, fácil de entender e rápido, “não foi muito burocrático”, sem um questionário grande.	Satisfação alta; o formulário curto foi visto como vantagem.
P3	Formulário e pós-teste	Leu o aviso de que a avaliação é anônima, notou o que outros alunos acharam e achou o fluxo “bem intuitivo”.	Satisfação alta.
P4	Revisão e pós-teste	Leu o termo de anonimato na revisão e achou “bem intuitivo”; no fim, disse que estava “super tranquilo”, que não havia “muito o que confundir” e que a avaliação estava “funcionando muito bem”.	Satisfação alta.
P5	Pós-teste	Disse que não tinha visto que dava para escolher o professor e que teria escolhido, porque a cobrança depende do professor; achou bom pesquisar e voltar para a página, “muito tranquilo”.	Campo de professor pouco percebido (AF-03); satisfação alta com a busca e a navegação.

Fonte: Victor Andreozzi de La Rocque Couto (2026), a partir da transcrição das gravações.

6 INTERPRETAÇÃO E ANÁLISE DOS DADOS

6.1 Análise de cada participante (intraparticipante)

Barbosa e Silva (2010, seção 9.7.4) recomendam interpretar primeiro os dados de cada participante, buscando respostas às perguntas do planejamento. Para cada sessão, a gravação da tela, as falas e as anotações foram analisadas em conjunto, como sugerem Rubin e Chisnell (2008, citados por Barbosa e Silva, 2010, seção 10.2.1), porque a análise conjunta revela o que um único tipo de dado não mostraria.

- **P1** explorou a interface antes de avaliar: abriu o filtro de professor "para ter noção do que tem", usou a busca com parte do nome e consultou as avaliações da disciplina antes de clicar em "Avaliar esta disciplina". No formulário, notou os comentários recentes ao lado e, na revisão, leu o termo de anonimato. Concluiu sem ajuda, conferiu que a avaliação já aparecia e avaliou outra disciplina sem comentário.
- **P2** leu a tela inicial com calma e chegou às avaliações pelo cartão "Avaliar disciplinas". Rolou a lista até a disciplina, notou as médias de dificuldade e cobrança, trocou o semestre sugerido e preferiu não informar o professor. Na revisão, leu o termo de anonimato e conferiu cada campo. Na confirmação, percebeu que o comentário não aparecia e só se tranquilizou ao encontrá-lo em "Ver avaliações".
- **P3** encontrou o cartão "Avaliar disciplinas" "no canto superior direito" da tela inicial, clicou em "Avaliar" no primeiro item da lista e leu o aviso de anonimato logo no início do formulário. Concluiu sem hesitar e achou o fluxo intuitivo.
- **P4** fez perguntas sobre o teste antes de começar (se avaliaria só a funcionalidade ou o sistema todo e se todas as disciplinas da UnB estavam no protótipo). Durante as tarefas, não pediu ajuda: buscou a disciplina, consultou as avaliações, leu o termo de anonimato na revisão, avaliou uma segunda disciplina "seguindo o mesmo caminho" e descartou a terceira lendo em voz alta a mensagem de que nada foi enviado.
- **P5** usou a busca nas três tarefas, mesmo quando o cenário pedia para avaliar "sem fazer nenhuma busca", o que indica que a busca é o caminho que ela prefere. Manteve "Prefiro não informar" no professor nas duas avaliações e o semestre sugerido (2026.1), e confirmou a revisão sem lê-la. Na Tarefa 3, hesitou sobre como sair, mas encontrou "Descartar" sozinha. No pós-teste, contou que não tinha visto que dava para escolher o professor.

6.2 Consolidação entre participantes (interparticipante)

Na consolidação, buscam-se recorrências, que distinguem características do grupo das idiossincrasias de cada participante (Barbosa e Silva, 2010, seção 9.7.5). As recorrências encontradas foram:

- **Todos concluíram as tarefas sem ajuda e sem desvios.** Os cinco encontraram a funcionalidade pela tela inicial, chegaram às disciplinas por caminhos diferentes (busca em 3 de 5, rolagem e primeiro item da lista) e usaram “Revisar e enviar” e “Confirmar envio” sem hesitar.
- **O fluxo é aprendido rápido.** Nas segundas avaliações, o tempo caiu para menos da metade em P1 e P4 e em um terço em P5.
- **O anonimato foi percebido por quatro dos cinco.** P3 leu o aviso do formulário e P1, P2 e P4 leram o termo da revisão, sem que o avaliador perguntasse. P5 não comentou o anonimato e passou pela revisão em menos de 4 segundos.
- **A informação dos outros alunos foi usada.** P1 e P4 abriram “Ver avaliações” antes de avaliar, P1 e P3 comentaram o resumo ao lado do formulário e P2 comentou as médias na lista.
- **Os valores já preenchidos tendem a ficar como estão (AF-02 e AF-03).** O semestre vem com 2026.1 e o professor com “Prefiro não informar”. P3 e P5 mantiveram o semestre sugerido, e P2 e P5 mantiveram o professor sem informação; P5 afirmou que não viu o campo e que teria escolhido o professor.
- **O descarte funcionou.** Os dois participantes que fizeram a Tarefa 3 saíram sem enviar, confirmando no diálogo de descarte, e nenhum envio acidental ocorreu.

Como idiossincrasias, P1 foi o único a explorar os filtros, P2 o único a ler cada item da revisão em voz alta e P2 o único a sentir falta do comentário na confirmação (AF-01). Mesmo não recorrentes, esses comportamentos são registrados porque podem indicar problemas que cinco participantes não bastam para confirmar.

6.3 Respostas às perguntas da avaliação

Tabela 9: Metas e resultados por pergunta

Nº	Pergunta	Meta aceitável / ideal	Resultado medido	Interpretação
1	O estudante conclui a avaliação sem ajuda?	3 de 5 / 5 de 5	5 de 5	Meta ideal atingida.
2	Quanto tempo leva para avaliar uma disciplina?	Até 3 min / até 2 min	Média de 1 min 42 s (± 18 s); 4 de 5 abaixo de 2 min e 5 de 5 abaixo de 3 min	Meta ideal atingida pela média, mesmo com o tempo das falas; P2 ficou na meta aceitável porque leu a tela inicial e a revisão em voz alta.
3	Ele encontra a disciplina pela busca e pelos filtros sem desvios?	Até 1 desvio / nenhum	0 desvios em todas as tarefas; 3 de 5 usaram a busca, com parte do nome (“int”, “comuni”, “psci”)	Meta ideal atingida. A busca por parte do nome funcionou em todas as tentativas.
4	Ele percebe, antes de enviar, que a avaliação é anônima?	4 de 5 / 5 de 5	4 de 5 leram o aviso ou o termo espontaneamente	Meta aceitável atingida; PP-02 resolvido. P5 confirmou a revisão sem lê-la.
5	Ele interpreta corretamente as escalas e seus extremos?	4 de 5 / 5 de 5	5 de 5 sem dúvidas; P3, P4 e P5 associaram as notas aos rótulos (“fácil”, “média”, “alta”)	Meta ideal atingida, com base nas falas; como as notas do teste são hipotéticas, não se avaliou se combinam com a disciplina real.
6	Ele consulta os relatórios enquanto preenche a avaliação?	3 de 5 / 4 de 5	3 de 5 (P1, P3 e P4) consultaram as avaliações ou o resumo ao lado	Meta aceitável atingida. Quem consultou o fez antes de avaliar, por iniciativa própria.
7	Ele desiste de uma avaliação sem enviá-la e sem perder o controle?	4 de 5 / 5 de 5	2 de 2 (P4 e P5) descartaram sem enviar e sem ajuda	Resultado ideal nos dois casos testados; P5 hesitou por alguns segundos sobre como sair.
8	Quão satisfeito ele fica com a experiência de avaliar?	4 de 5 / 5 de 5	5 de 5 avaliaram a experiência de forma positiva (Tabela 8)	Meta ideal atingida (“bem intuitivo”, “não foi burocrático”, “muito tranquilo”); a única crítica, de P5, virou o problema AF-03.

Fonte: Victor Andreozzi de La Rocque Couto (2026), a partir das metas da Tabela 11 do planejamento.

6.4 Confiabilidade e validade

Os resultados de uma avaliação de IHC indicam tendências de problemas, e não a certeza de que vão ocorrer no uso real; da mesma forma, não encontrar um problema não prova que ele não exista, apenas que o estudo não o revelou no escopo avaliado (Barbosa e Silva, 2010, seção 9.7.5). Neste teste, os limites são:

- cinco participantes revelam a maior parte dos problemas, mas as médias de tempo não têm significância estatística, e as Tarefas 2 e 3 foram feitas por apenas dois participantes;

- o pensar em voz alta aumentou os tempos e pode ter levado os participantes a ler a interface com mais atenção do que no uso real (Barbosa e Silva, 2010, seção 9.7.3);
- o avaliador estava ao lado do participante, e os participantes eram estudantes próximos do contexto do avaliador, o que pode tê-los deixado menos críticos (Barbosa e Silva, 2010, seção 9.7.2);
- os dados das disciplinas são fictícios.

Por outro lado, a validade ecológica é maior que a do protótipo de papel: cada participante usou a interface sozinho, com rolagem, digitação e estados reais, e avaliou disciplinas que de fato cursou.

7 PROBLEMAS ENCONTRADOS

Como pedem Barbosa e Silva (2010, seção 10.2.1), cada problema é apresentado com o local, a descrição e a justificativa, a discussão dos fatores de usabilidade prejudicados e a sugestão de solução. Seguindo Kuniavsky (2003, citado por Barbosa e Silva, 2010, seção 10.2.1), os problemas foram agrupados por categoria e acompanhados de uma hipótese sobre a causa. A severidade segue a escala de 0 a 3 usada nas avaliações anteriores.

Tabela 10: Problemas encontrados no teste de usabilidade

ID	Local	Descrição e justificativa	Discussão: fatores de usabilidade prejudicados	Severidade (0 a 3)	Frequência	Sugestão de solução
Categoria: valores já preenchidos no formulário						
AF-03	T4 Formulário, campo “Professor (opcional)”	O campo vem com “Prefiro não informar”. P5 manteve esse valor nas duas avaliações e disse no pós-teste que não tinha visto que dava para escolher o professor, e que teria escolhido, porque “depois você avalia se ele cobra ou não”. Hipótese: o rótulo “opcional” e o valor já preenchido fazem o campo parecer resolvido, e ele fica longe da escala de cobrança, que depende dele.	Eficácia e utilidade: uma nota de “cobrança do professor” sem professor perde sentido para quem consulta os relatórios por professor, que é o uso principal da funcionalidade.	2 Grande	1 de 5 não percebeu o campo; 2 de 5 mantiveram o valor padrão	Deixar o campo sem valor, exigir uma escolha (com “Prefiro não informar” como opção explícita) e repetir o nome do professor escolhido no título da escala de cobrança.
AF-02	T4 Formulário, campo “Semestre em que cursou”	O campo vem com o semestre atual (2026.1). P1, P2 e P4 trocaram para o semestre em que cursaram a disciplina; P3 e P5 mantiveram 2026.1. Hipótese: um valor padrão parece já respondido, e o estudante pode não revisá-lo.	Eficácia e segurança no uso: a avaliação pode ser publicada com o semestre errado, o que distorce as médias por semestre e por professor.	1 Leve	3 de 5 precisaram trocar; 2 de 5 mantiveram o padrão	Deixar o campo sem valor e torná-lo obrigatório, ou sugerir o último semestre concluído e destacar o campo na revisão.
Categoria: retorno do sistema depois do envio						

ID	Local	Descrição e justificativa	Discussão: fatores de usabilidade prejudicados	Severidade (0 a 3)	Frequência	Sugestão de solução
AF-01	T6 Confirmação	O resumo da confirmação mostra dificuldade, cobrança, semestre e "Anônima", mas não o comentário. P2 disse que "só não apareceu aqui o comentário" e só se tranquilizou ao encontrá-lo em "Ver avaliações"; P1 também foi conferir a avaliação publicada. Hipótese: como o resto da avaliação aparece no resumo, a ausência do comentário é lida como possível falha no envio.	Segurança no uso e satisfação: o participante fica em dúvida se a ação foi concluída e precisa conferir em outra tela (visibilidade do estado do sistema).	1 Leve	1 de 5 (P2); 2 de 5 conferiram a publicação	Incluir o comentário, ou um trecho dele, no resumo da confirmação; quando não houver comentário, mostrar "Sem comentário".
Categoria: atenção à revisão						
AF-04	T5 Revisão	P5 confirmou as duas avaliações em menos de 4 segundos, sem ler a revisão nem o termo de anonimato, e foi a única que não percebeu o anonimato. Hipótese: depois do formulário, o botão "Confirmar envio" fica em destaque e a revisão é vista como uma etapa a pular.	Segurança no uso: a revisão, criada para corrigir PP-03, só protege quem a lê; os erros de AF-02 e AF-03 passariam por ela sem ser notados.	1 Leve	1 de 5	Destacar na revisão os campos com valor padrão (semestre e professor), com um aviso do tipo "Confira o semestre e o professor".

Fonte: Victor Andreozzi de La Rocque Couto (2026). Severidade conforme Barbosa e Silva (2010): 0 Cosmético, 1 Leve, 2 Grande, 3 Catastrófico. Problemas ordenados pela prioridade (severidade e frequência).

Nenhum problema catastrófico foi observado, e nenhum impediu a conclusão de uma tarefa. Esse resultado é coerente com a Entrega 6: os problemas mais graves do protótipo de papel foram corrigidos antes do teste, e o teste interno do protótipo HTML corrigiu outros seis. Os quatro problemas encontrados afetam a qualidade dos dados registrados (semestre e professor) e a confiança no envio, e não a capacidade de avaliar.

Tabela 11: Situação dos problemas das avaliações anteriores

Problema anterior	Correção no protótipo	Situação no teste
PP-01 Escala de estrelas lida como nota geral	Escalas de 1 a 5 com rótulos nos extremos	Resolvido: nenhum participante teve dúvida sobre as escalas.
PP-02 Sem aviso de anonimato	Aviso no formulário e termo na revisão	Resolvido: 4 de 5 leram o aviso ou o termo; ver AF-04.
PP-03 Sem revisão nem confirmação	Tela de revisão e tela de confirmação	Resolvido: 5 de 5 passaram pela revisão e pela confirmação; surgiram os ajustes AF-01 e AF-04.
PP-04 "Ver disciplina" lido como ementa	"Ver avaliações" e médias nos resultados	Resolvido: P1 e P4 usaram "Ver avaliações" com o sentido esperado, e P2 leu as médias na lista.
PP-05 Sem "Descartar"	Botão "Descartar" na barra inferior e diálogo de confirmação	Resolvido: P4 e P5 descartaram sem enviar e leram que nada foi enviado.
PP-06 Comentário lido como sugestão ao curso	Título "Comentário para os próximos alunos" e texto de apoio	Resolvido: os comentários foram dirigidos a outros alunos ("a matéria é bastante reflexiva", "Professor atencioso").
B-01 Sem caminho para descartar o rascunho	Botão "Descartar" sempre visível e diálogo de confirmação	Resolvido: P4 e P5 descartaram sem enviar (ver também PP-05).
B-02 Botão de confirmar fora da área visível e sem indicação de processamento	"Revisar e enviar" e "Confirmar envio" na barra inferior e estado "Enviando…"	Resolvido: ninguém precisou procurar o botão de envio, e o estado "Enviando…" apareceu sem causar dúvida.
A-01 a A-03 e B-03 a B-07	Ver Tabela 10 do planejamento	Nenhum reapareceu.

Fonte: Victor Andreozzi de La Rocque Couto (2026), a partir do relato do protótipo de papel (Tabela 7) e do planejamento da avaliação (Tabela 10).

8 PLANEJAMENTO PARA O REPROJETO

O relato termina com um planejamento para o reprojeto do sistema (Barbosa e Silva, 2010, seção 9.7.5). As correções foram priorizadas pela severidade e pela frequência, e as explicações qualitativas de cada problema justificam as mudanças propostas, como recomendam Rubin e Chisnell (2008, citados por Barbosa e Silva, 2010, seção 10.2.1).

Tabela 12: Prioridades do reprojeto

Prioridade	Problema	Tela	Mudança proposta	Como verificar
1	AF-03	T4 Formulário	Professor sem valor padrão e obrigatório, com “Prefiro não informar” como escolha explícita; nome do professor no título da escala de cobrança.	Observar se os participantes escolhem o professor sem que o avaliador mencione o campo.
2	AF-02	T4 Formulário	Semestre sem valor padrão e obrigatório, com a mensagem de validação junto ao campo.	Conferir se o semestre registrado corresponde ao que o participante diz ter cursado.
3	AF-04	T5 Revisão	Destacar o semestre e o professor na revisão, com o aviso “Confira o semestre e o professor”.	Medir o tempo na revisão e se algum participante volta para editar.
4	AF-01	T6 Confirmação	Mostrar o comentário (ou um trecho) no resumo, ou “Sem comentário”.	Observar se o participante procura o comentário depois de confirmar.

Fonte: Victor Andreozzi de La Rocque Couto (2026).

9 CONCLUSÃO

O protótipo de alta fidelidade da Avaliação de Disciplinas atingiu as metas de eficácia e eficiência: os cinco participantes avaliaram uma disciplina sem ajuda e sem desvios, em média em 1 min 42 s, e os que repetiram a tarefa levaram menos da metade do tempo. As correções do protótipo de papel funcionaram, em especial o aviso de anonimato, a revisão, o "Ver avaliações" e o descarte. Os problemas que restam são leves ou grandes, mas não impedem o uso: os valores já preenchidos de semestre e professor podem gerar avaliações com dados errados, o que ameaça justamente a utilidade dos relatórios para quem monta a grade.

Como o design de IHC é iterativo, recomenda-se aplicar as mudanças da Tabela 12 e fazer uma rodada curta de testes antes da implementação, que verifique se as correções resolvem os problemas sem criar outros (Barbosa e Silva, 2010, seção 9.7.5).

10 CUIDADOS ÉTICOS

Os cinco participantes foram informados dos objetivos e do procedimento, viram o termo de consentimento e autorizaram em voz alta, na gravação, o uso de imagem e áudio, cientes de que podiam sair a qualquer momento (Barbosa e Silva, 2010, seção 5.4). Neste relato, eles são identificados apenas por códigos, e nomes ditos durante as sessões foram omitidos. Como as gravações mostram o rosto, elas não são publicadas no site e ficam com acesso restrito ao avaliador.

11 DECLARAÇÃO SOBRE O USO DE IA GENERATIVA

A ferramenta de IA Generativa Claude (Anthropic) foi usada para transcrever o áudio das cinco sessões (com o modelo Whisper, executado localmente), medir os tempos pelas mudanças de tela das gravações, redigir o relato a partir das gravações, das transcrições, dos planejamentos da Entrega 6 e das obras de referência, e montar este documento. O conteúdo, as decisões e a revisão final são responsabilidade do autor, Victor Andreozzi de La Rocque Couto.

12 REFERÊNCIAS

BARBOSA, Simone Diniz Junqueira; SILVA, Bruno Santana da. **Interação Humano-Computador**. Rio de Janeiro: Elsevier / Campus, 2010.

PREECE, Jennifer; ROGERS, Yvonne; SHARP, Helen. **Design de Interação**: além da interação homem-computador. Porto Alegre: Bookman, 2005.

SANTA ROSA, José Guilherme; MORAES, Anamaria de. **Avaliação e projeto no design de interfaces**. Teresópolis: 2AB, 2008.