



instituto superior de
engenharia de lisboa

Instituto Superior de Engenharia de Lisboa
Licenciatura Engenharia Informática e Multimédia

Ano letivo 2025/2026

Interação Pessoa-Máquina
Trabalho Prático nº3

Turma 52D

Miguel Azenha, A47708
Maria Luísa Sampaio, A50799

Dezembro de 2025

Índice

1	Introdução	1
1.1	Definição do Problema	1
1.2	Objetivos da Avaliação	1
1.3	Sistemas em Comparaçāo e Metodologia	2
2	Metodologia de Avaliação	3
2.1	Participantes e Ética	3
2.2	Desenho Experimental	3
2.3	Variável Independente	3
2.4	Variáveis Dependentes	3
2.5	System Usability Scale (SUS)	4
2.5.1	Estrutura e Escala	4
2.5.2	Cálculo da Pontuação	4
2.6	Tarefas e Procedimento	5
2.6.1	Procedimento	5
2.6.2	Métricas	5
3	Perfil dos Participantes	6
3.1	Critérios de Amostragem e Recrutamento	6
3.2	Caracterização da Amostra	6
4	Análise de Resultados	8
4.1	Análise de Descritiva	8
4.1.1	Satisfação SUS	8
4.1.2	Teste de Hipótese	9
4.1.3	Análise de Eficiência (Tempo)	11
4.1.4	Análise de Eficácia (Erros)	12
5	Avaliações	13
6	Conclusões	14

Índice de figuras

1	Gráfico sobre a frequencia dos usos de transportes	6
2	Gráfico sobre as idades dos participantes	7
3	Gráfico de Comparação de Tempos Médios em Cada Tarefa	11

1 Introdução

A presente avaliação experimental e formal é realizada no âmbito da Unidade Curricular de Interação Pessoa–Máquina e constitui o 3º Trabalho Prático. Este trabalho dá continuidade às fases de investigação e design centradas no utilizador executadas nos trabalhos práticos anteriores.

1.1 Definição do Problema

O foco central deste trabalho insere-se na análise e superação da fricção induzida pela dependência do cartão físico do passe Navegante no quotidiano dos utilizadores de transportes públicos. A fase inicial de investigação identificou que a gestão do título de transporte físico cria múltiplos pontos de fricção no fluxo diário. Estes constrangimentos incluem:

1. **Esquecimento:** A dependência do cartão físico é uma fonte significativa de frustração, levando a situações recorrentes de esquecimento do passe em casa.
2. **Ineficiência na Validação:** A necessidade de procurar o cartão na mala ou no bolso resulta em tempos médios de procura observados entre 12 a 35 segundos, ou em alguns casos superiores caso o cartão esteja dentro de uma mochila ou mala. Esta fricção causa filas, stress e pequenos atrasos acumulados na entrada dos transportes, especialmente nos autocarros. Adicionalmente, os validadores nos autocarros falham frequentemente a leitura à primeira tentativa, exigindo que o utilizador volte a tentar.
3. **Complexidade no Carregamento:** O processo de renovação do passe, efetuado maioritariamente em máquinas automáticas, Multibanco ou espaços Navegante, gera filas de espera e frustração, especialmente no início do mês ou em casos de avaria das máquinas.
4. **Falta de Integração Digital:** Existe frustração por parte dos utilizadores devido à ausência de uma versão digital completa do passe e à falta de integração com sistemas de carteira móvel amplamente utilizados, como a *Apple Wallet* ou *Google Wallet*.

Estes fatores de frustração e ineficiência representam um desafio significativo no fluxo quotidiano, com os participantes a valorizarem soluções digitais que eliminem a dependência do cartão físico e que permitam a validação direta através do telemóvel. A solução proposta consiste num protótipo de alta fidelidade que visa digitalizar o passe e integrar funcionalidades essenciais, como carregamento automático e notificações de saldo baixo.

1.2 Objetivos da Avaliação

O presente estudo visa realizar uma avaliação rigorosa do protótipo desenvolvido. A metodologia adotada segue o formato de uma avaliação experimental e formal, comparando o protótipo com um sistema semelhante que atua como *baseline*.

Os objetivos específicos desta avaliação são:

1. **Avaliar a Usabilidade e a Experiência do Utilizador:** Mensurar o nível de eficácia, eficiência e satisfação, recorrendo a métricas ordinais e contínuas, durante a execução de tarefas-chave no protótipo de alta fidelidade.
2. **Comparação com o Sistema *Baseline*:** Determinar se o protótipo digital oferece uma melhoria estatisticamente significativa no desempenho do utilizador em comparação com a solução de referência existente ou com um sistema semelhante que realize a mesma tarefa.

3. **Validação dos Requisitos de Design:** Verificar se as funcionalidades implementadas, como a integração digital do passe ou o fluxo de carregamento, satisfazem os requisitos identificados durante a fase de investigação.

1.3 Sistemas em Comparaçāo e Metodologia

O sistema avaliado consiste num protótipo de alta fidelidade de uma aplicação móvel para gestão e validação do passe Navegante, desenvolvido em Figma. Este sistema será comparado com um sistema de controlo que permita realizar a mesma tarefa principal de gestão e validação de títulos de transporte.

A avaliação será conduzida através de uma metodologia experimental rigorosa, incluindo a definição de variáveis dependentes (métricas ordinais e contínuas), o recrutamento de 10 a 16 participantes e a análise estatística dos resultados, conforme detalhado no capítulo seguinte.

2 Metodologia de Avaliação

A metodologia de avaliação foi delineada para conduzir um estudo experimental e formal, conforme as diretrizes do enunciado do trabalho prático. O principal objetivo foi medir e comparar o desempenho dos utilizadores e a satisfação subjetiva ao executar tarefas essenciais no protótipo de alta fidelidade desenvolvido em comparação com um sistema *baseline* já existente.

2.1 Participantes e Ética

O estudo incluiu um total de 10 participantes, satisfazendo o requisito de ter entre 10 e 16 indivíduos. O perfil dos participantes foi selecionado para ser representativo dos utilizadores-alvo do sistema, consistindo em utilizadores frequentes de transportes públicos em Lisboa que possuem o Passe Navegante.

Todos os participantes foram informados sobre os objetivos e o procedimento da avaliação e confirmaram a sua participação de forma voluntária, assinando um formulário de consentimento informado. O tratamento dos dados seguiu princípios éticos, garantindo a confidencialidade das respostas. O perfil detalhado de cada participante (idade, frequência de uso de transportes, etc.) será reportado no capítulo de Resultados.

2.2 Desenho Experimental

Foi adotado um *Desenho Experimental Intra-Sujeitos (Within-Subjects)*. Neste desenho, cada participante foi exposto a ambas as condições de teste (Protótipo vs. Controlo), o que permitiu que o próprio indivíduo atuasse como seu controlo. Este desenho é inherentemente mais eficiente do ponto de vista estatístico, uma vez que controla as diferenças de variabilidade interindividual.

Para mitigar o efeito de aprendizagem, que ocorre quando a exposição a uma condição afeta o desempenho na condição seguinte, o estudo utilizou a técnica de balanceamento através da rotação das condições. Metade dos participantes iniciou as tarefas com o Protótipo e depois passou para o sistema *baseline*, enquanto a outra metade seguiu a ordem inversa.

2.3 Variável Independente

A *Variável Independente* (o fator manipulado) foi o Sistema em Avaliação, com dois níveis:

1. **Condição A (Protótipo)**: O protótipo de alta fidelidade da aplicação Navegante (T2).
2. **Condição B (Controlo/Baseline)**: Um sistema semelhante que permite a realização da mesma tarefa de gestão/validação, simulando o processo atual (ex: a aplicação existente Navegante ou o uso do cartão físico).

2.4 Variáveis Dependentes

As *Variáveis Dependentes* (as métricas de avaliação) foram selecionadas para medir a eficácia, a eficiência e a satisfação do utilizador, e incluíram obrigatoriamente dados ordinais (escalas de Likert) e dados intervalares contínuos:

1. Métricas Contínuas

- **Tempo de Execução (Task Completion Time)**: Tempo em segundos gasto para completar cada tarefa. Estas medições foram efetuadas manualmente pelo avaliador, cronometrando o tempo desde o início da tarefa até à sua conclusão.

- **Taxa de Erros:** Contagem do número de erros cometidos pelos participantes durante a execução de cada tarefa. O número de erros é calculado pelo número de cliques feitos pelo participante a subtrair pelo número mínimo de cliques para completar a tarefa.

2. Métrica Ordinal

- **Satisfação Subjetiva :** Avaliada através da System Usability Scale (SUS).

2.5 System Usability Scale (SUS)

O *System Usability Scale* (SUS) é um instrumento amplamente utilizado para a avaliação da usabilidade de uma vasta gama de produtos e interfaces de utilizador .

O principal valor do SUS consiste em fornecer um único objetivo de referência que representa a percepção subjetiva dos participantes relativamente à usabilidade de um produto ou serviço. É considerado uma ferramenta económica e eficaz, aplicável a uma vasta gama de tecnologias, sendo independente da tecnologia avaliada.

2.5.1 Estrutura e Escala

O SUS é um questionário auto-administrado composto por 10 afirmações, classificadas numa escala de Likert de 5 pontos (de 1 a 5).

As afirmações alternam entre formulações positivas e negativas. Esta alternância foi implementada para reduzir o viés de resposta e garantir que os inquiridos leem cada afirmação de forma cuidadosa.

O questionário pode ser dividido em duas subescalas:

- **Usable:** composta por 8 itens;
- **Learnable (Aprendizagem):** composta pelos itens 4 e 10.

2.5.2 Cálculo da Pontuação

A pontuação final do SUS varia entre 0 e 100. O método de cálculo da pontuação é descrito da seguinte forma:

- Para os itens ímpares (1, 3, 5, 7 e 9), a contribuição da pontuação corresponde ao valor selecionado na escala menos 1.
- Para os itens pares (2, 4, 6, 8 e 10), a contribuição corresponde a 5 menos o valor selecionado na escala.
- A contribuição individual de cada item varia entre 0 e 4.
- A soma das contribuições de todos os itens é multiplicada por 2.5, resultando na pontuação final do SUS.

Uma pontuação SUS superior a 68 é geralmente interpretada como estando acima da média em termos de usabilidade.

2.6 Tarefas e Procedimento

As tarefas foram selecionadas para cobrir os pontos de fricção identificados, focando-se na aquisição e validação de títulos de transporte:

- **Tarefa 1: Carregar o Passe:** Simular a aquisição ou recarregamento de um título de transporte mensal.
- **Tarefa 2: Validar o Passe:** Simular o processo de validação de uma viagem.
- **Tarefa 3: Ver Perfil/Informação do Passe:** Simular a consulta de informações pessoais e do estado do cartão.

2.6.1 Procedimento

A avaliação foi realizada em sessões individuais, seguindo os seguintes passos:

1. **Introdução:** O participante assinou o consentimento informado, preencheu um breve questionário de perfil e recebeu instruções gerais sobre o teste, incluindo a explicação do método *pensar alto* (verbalização de pensamentos e sentimentos durante a interação).
2. **Execução das Tarefas:** O participante executou as três tarefas sequencialmente.
3. **Recolha de Dados Contínuos:** O avaliador registou manualmente o Tempo de Execução e a Taxa de Erros para cada tarefa e condição.
4. **Recolha de Dados Ordinais:** Após a conclusão de todas as tarefas em cada sistema, o participante preencheu o questionário SUS relativo à sua experiência com essa condição.
5. **Conclusão:** Os participantes foram sumariamente informados sobre os objetivos específicos do estudo (*debriefing*).

2.6.2 Métricas

Dado que as tarefas 2 e 3 correspondem a funcionalidades inovadoras propostas pelo protótipo e inexistentes na aplicação atual, adotou-se o seguinte critério de mediação para o sistema de controlo:

- **Tempo de Execução:** Foi contabilizado o tempo que cada participante demorou na procura ativa da funcionalidade até verbalizar a sua inexistência ou desistir da tarefa.
- **Taxa de Erros:** Foi considerada **nula** para efeitos de comparação direta, pois qualquer interação seria exploratória e não um erro de execução.

3 Perfil dos Participantes

Com base nas informações recolhidas ao longo do estudo, foi elaborado o perfil dos participantes, que inclui a descrição da amostra (características demográficas e padrões de uso).

3.1 Critérios de Amostragem e Recrutamento

Os participantes foram recrutados com base em critérios de inclusão que visavam garantir a relevância da amostra para o contexto do estudo, nomeadamente serem utilizadores habituais do passe Navegante e demonstrarem motivação para cooperar e participar voluntariamente na experiência.

A recolha de dados foi realizada mediante o preenchimento de um formulário de consentimento informado (anexado no final deste documento) por todos os participantes. Este procedimento ético assegurou a participação voluntária e garantiu que os indivíduos tinham conhecimento prévio dos objetivos e dos procedimentos do estudo.

3.2 Caracterização da Amostra

Tabela 1: Caracterização geral da amostra

Característica	Descrição
Tamanho da Amostra	N = 10 (total de participantes)
Frequência de Uso	Maioria dos participantes eram utilizadores frequentes do passe Navegante.
Faixa Etária	Idades compreendidas entre os 17 e os 48 anos.
Uso da Aplicação Original	A percepção geral dos participantes indicou que a adoção da aplicação oficial <i>Navegante</i> é reduzida, embora esta seja considerada funcional.

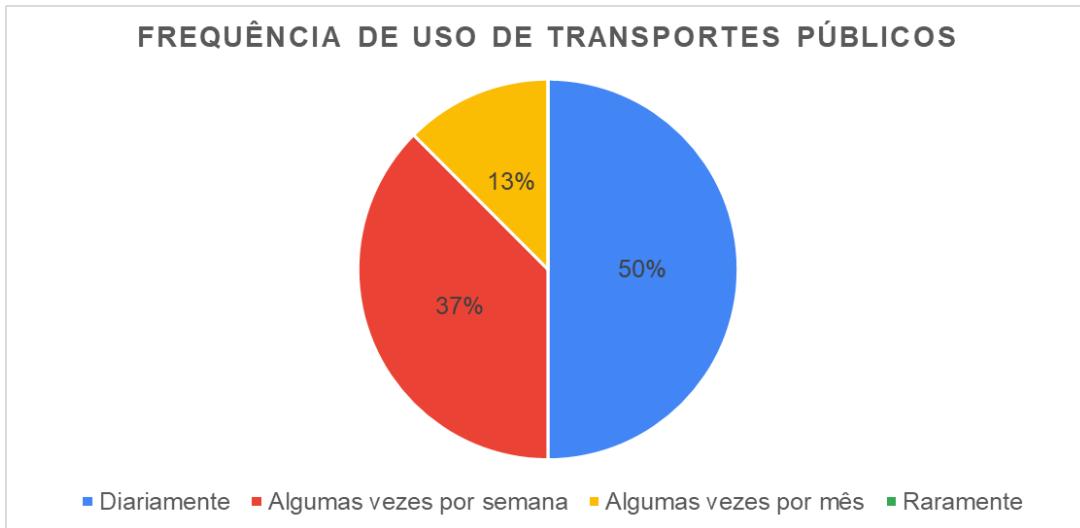


Figura 1: Gráfico sobre a frequencia dos usos de transportes

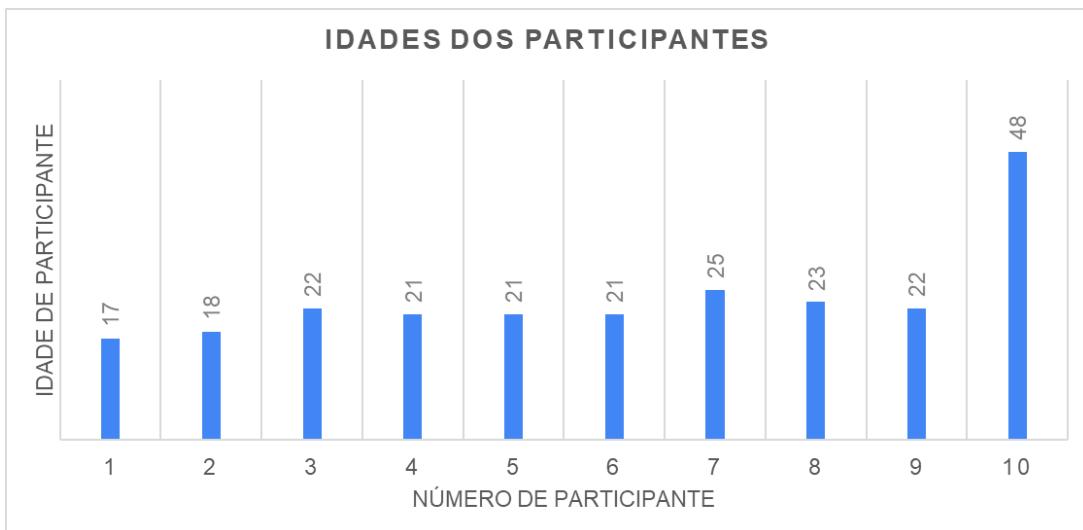


Figura 2: Gráfico sobre as idades dos participantes

4 Análise de Resultados

4.1 Análise de Descritiva

4.1.1 Satisfação SUS

As perguntas feitas seguiram o seguinte formato: [martins2015european]

Fórmula para o SUS Para cada um dos 10 participantes, calcula-se uma pontuação individual:

1. **Perguntas Ímpares:** Subtrair 1 à resposta do utilizador. ($R - 1$)
2. **Perguntas Pares:** Subtrair à resposta do utilizador 5. ($5 - R$)
3. **Soma Total:** Somar os resultados dos passos 1 e 2.
4. **Multiplicador:** Multiplicar a soma por 2.5 para obter a escala de 0-100.

Os participantes responderam as mesmas perguntas tanto para a aplicação atual como para o protótipo demonstrado.

Resultados App Atual

Participante nº1 Respostas: 3, 2, 2, 1, 3, 3, 5, 1, 4, 1

Perguntas Ímpares

$$\begin{aligned}3 - 1 &= 2 \\2 - 1 &= 1 \\3 - 1 &= 2 \\5 - 1 &= 4 \\4 - 1 &= 3 \\\Sigma &= 2 + 1 + 2 + 4 + 3 = 12\end{aligned}$$

Perguntas Pares

$$\begin{aligned}5 - 2 &= 3 \\5 - 1 &= 4 \\5 - 3 &= 2 \\5 - 1 &= 4 \\5 - 1 &= 4 \\\Sigma &= 3 + 4 + 2 + 4 + 4 = 17\end{aligned}$$

Cálculo Final

- Soma Total = $12 + 17 = 29$
- Pontuação SUS = $29 * 2.5 = 72.5$

Fazendo estes cálculos para todos os participantes, obtiveram-se os seguintes resultados:

Part.	Ímpares	Pares	Soma	SUS
1	12	17	29	72,5
2	12	17	29	72,5
3	13	20	33	82,5
4	15	20	35	87,5
5	14	17	31	77,5
6	16	19	35	87,5
7	14	17	31	77,5
8	4	7	11	27,5
9	9	11	20	50
10	10	9	19	47,5

Tabela 2: Tabela de Pontuação dos Participantes para a Aplicação Atual

$$\text{Média: } M = \frac{\Sigma x}{N} = \frac{682,5}{10} = 68,5$$

Resultados Protótipo

Seguindo os mesmos cálculos feitos acima:

Part.	Ímpares	Pares	Soma	SUS
1	18	18	36	90
2	20	18	38	95
3	18	19	37	92,5
4	19	20	39	97,5
5	18	20	38	95
6	20	20	40	100
7	16	18	34	85
8	14	16	30	75
9	17	16	33	82,5
10	16	15	31	77,5

Tabela 3: Tabela de Pontuação dos Participantes para a Aplicação Atual

$$\text{Média: } M = \frac{\Sigma x}{N} = \frac{890}{10} = 89$$

4.1.2 Teste de Hipótese

Para provar estatisticamente que a melhoria não foi "sorte", calcula-se o **Paired T-Test**

Fórmula

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$

Onde:

- \bar{d} = Média das diferenças entre os dois sistemas
- s_d = Desvio padrão das diferenças.
- n = Número de Participantes (10).

Começa-se por calcular a diferença SUS para cada participante:

Part.	App Atual	Protótipo	Diferença
1	72,5	90	+17,5
2	72,5	95	+22,5
3	82,5	92,5	+10,0
4	87,5	97,5	+10,0
5	77,5	95	+17,5
6	87,5	100	+12,5
7	77,5	85	+7,5
8	27,5	75	+47,5
9	50	82,5	+32,5
10	47,5	77,5	+30,0
Média	68,3	89,0	+20,75

Tabela 4: Comparaçāo de pontuação SUS entre a Aplicação Atual e o Protótipo

De seguida, calcula-se o Desvio Padrão (s_d):

$$s_d = \sqrt{\frac{\sum(d_i - \bar{d})^2}{n - 1}}$$

$$\begin{aligned}
 P1: (17,5 - 20,75)^2 &= (-3,25)^2 &= 10,56 \\
 P2: (22,5 - 20,75)^2 &= (1,75)^2 &= 3,06 \\
 P3: (10,0 - 20,75)^2 &= (-10,75)^2 &= 115,56 \\
 P4: (10,0 - 20,75)^2 &= (-10,75)^2 &= 115,56 \\
 P5: (17,5 - 20,75)^2 &= (-3,25)^2 &= 10,56 \\
 P6: (12,5 - 20,75)^2 &= (-8,25)^2 &= 68,06 \\
 P7: (7,5 - 20,75)^2 &= (-13,25)^2 &= 175,56 \\
 P8: (47,5 - 20,75)^2 &= (26,75)^2 &= 715,56 \\
 P9: (32,5 - 20,75)^2 &= (-11,75)^2 &= 138,06 \\
 P10: (30,0 - 20,75)^2 &= (9,25)^2 &= 85,56 \\
 \Sigma &= 1438,125
 \end{aligned}$$

$$\text{Variância}(s^2) = \frac{1438,125}{10 - 1} \approx 159,79$$

$$s_d = \sqrt{157,79} \approx 12,64$$

O desvio padrão de 12,64 indica que existe uma variabilidade considerável na forma como as pessoas reagiram à mudança.

Depois, calcula-se o Erro Padrão (SE):

$$SE = \frac{s_d}{\sqrt{n}} = \frac{12,64}{\sqrt{10}} = \frac{12,64}{3,162} \approx 4,00$$

Calcula-se agora o Valor T ($t\text{-value}$)

$$t = \frac{\bar{d}}{SE} = \frac{20,75}{4,00} \approx 5,19$$

Por fim, calcula-se o valor P ($p\text{-value}$)

Para encontrar o $p\text{-value}$, tem que se consultar a Tabela de Distribuição T com:

1. **Graus de Liberdade (df)**: $n - 1 = 9$
2. **Valor T**: 5,19

Analizando na tabela T para $df = 9$:

- Para $p = 0.05$, o t crítico é **1,833**.
- Para $p = 0.001$, o t crítico é **4,297**

Como o valor t é superior a 4,297, o $p\text{-value}$ é menor do que 0.001.

4.1.3 Análise de Eficiência (Tempo)

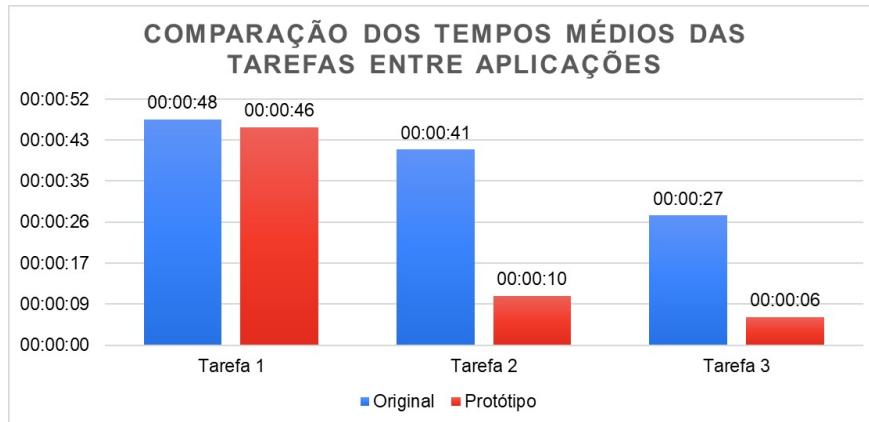


Figura 3: Gráfico de Comparação de Tempos Médios em Cada Tarefa

A figura 3 mostra uma comparação de tempos médios nas diferentes tarefas na aplicação atual e no protótipo.

A análise dos tempos de execução revelou 2 cenários distintos. Nas tarefas 2 e 3, o protótipo reduziu o tempo médio em cerca de 75%, comprovando a eficácia da nova arquitetura de informação.

Contudo, na Tarefa 1, não se registou uma melhoria significativa no tempo médio (47,6s na App Atual vs 45,9 no Protótipo). Esta estagnação deve-se a um problema identificado na interface do protótipo: o botão principal "Adicionar/Comprar" induziu os utilizadores a acreditar que iniciariam o processo de compra imediata, mas encaminhou-os para um ecrã intermédio de leitura de cartão. Este passo extra inesperado quebrou o fluxo mental do utilizador, gerando hesitações que anularam os ganhos de eficiência esperados.

4.1.4 Análise de Eficácia (Erros)

5 Avaliações

6 Conclusões

A realização deste segundo trabalho prático permitiu dar continuidade ao estudo iniciado na primeira fase, passando da análise do problema e das necessidades dos utilizadores para a geração de soluções concretas e a sua materialização através da ideação e da prototipagem de baixa fidelidade.

A solução proposta procura eliminar a dependência do cartão físico, simplificando tarefas quotidianas como o carregamento e a validação do passe. Este conceito demonstra não só um potencial de melhoria da experiência do utilizador, como também um alinhamento com as tendências atuais de mobilidade digital e integração tecnológica nos transportes públicos.

O protótipo de baixa fidelidade revelou-se uma ferramenta essencial para comunicar ideias, testar fluxos de interação e identificar pontos de melhoria sem necessidade de recorrer a soluções técnicas complexas. O seu caráter flexível e iterativo contribuiu para um design mais centrado no utilizador e preparado para evoluir com base em futuros testes.

A transição do protótipo de baixa para alta fidelidade permitiu transformar um conceito funcional num produto visualmente apelativo e tecnicamente coerente. As melhorias aplicadas reforçaram a intuitividade, eficiência e acessibilidade da aplicação, refletindo o contributo direto dos testes com utilizadores.

O protótipo de alta fidelidade constitui, assim, a base para um futuro desenvolvimento real da aplicação Navegante Digital, alinhado com as práticas de design centrado no utilizador.