



# Relatório dos resultados obtidos com o trabalho de avaliação laboratorial do protótipo da plataforma CeNTER

Deliverable D3.3 – Parte 1/2

<b>Referência do Projeto</b>	CENTRO-01-0145-FEDER-000002
<b>Data de início</b>	01/04/2017
<b>Duração</b>	36 meses
<b>Entrega n.º</b>	D3.3
<b>Workpackage</b>	WP3
<b>Nível de disseminação*</b>	PU
<b>Versão</b>	4.0
<b>Data</b>	29/04/2020
<b>Autor(es)</b>	Fernando Ramos (coord.); Ana Margarida Almeida; Maria João Antunes; Luís Pedro; Eliza Oliveira; Oksana Tymoshchuk; Daniel Carvalho.
<b>Email</b>	fernando.ramos@ua.pt (coord.)

\*Níveis de disseminação: PU= Público; CO= Confidencial

Este documento foi elaborado no âmbito do Programa Integrado de IC&DT "CeNTER – Redes e Comunidades para a Inovação Territorial" (CENTRO-01-0145-FEDER-000002), financiado pelo Programa Operacional Regional do Centro (CENTRO 2020), PT2020.

Promotor:



Cofinanciado por:



CENTRO-01-0145-FEDER-000002

## Histórico de Edições

V1.0	01-03-2020	Modificações a nível de estrutura, índice, nome dos títulos e conteúdo a ser desenvolvido nos respetivos tópicos	Fernando Ramos Maria João Antunes Luís Pedro Margarida Almeida Eliza Oliveira Oksana Tymoshchuk Daniel Carvalho
V2.0	15-04-2020	Modificações a nível de figuras, conteúdo textual e nomenclatura utilizadas ao longo do texto.	Fernando Ramos Maria João Antunes Luís Pedro Margarida Almeida Eliza Oliveira Oksana Tymoshchuk Daniel Carvalho
V3.0	30-04-2020	Correção final das modificações feitas e outras considerações.	Fernando Ramos Maria João Antunes Luís Pedro Margarida Almeida Eliza Oliveira Oksana Tymoshchuk Daniel Carvalho
V4	10-08-2020	Adição das informações sobre os testes de Acessibilidade e Microcopy	Eliza Oliveira

## Contribuições

- Eliza Oliveira
- Oksana Tymoshchuk
- André Castelo Branco
- Daniel Carvalho
- Fernando Ramos
- Maria João Antunes
- Luís Pedro
- Margarida Almeida

## Índice

1.	Introdução	4
2.	Design Centrado no Utilizador	4
3.	Avaliação da User Experience	6
4.	Teste com Peritos	8
4.4.	Objetivos dos Testes	8
4.5.	Metodologia	9
4.6.	Perfil dos Avaliadores	10
4.7.	Ambiente dos Testes	11
4.8.	Versões do Protótipo	11
5.	Resultados dos Testes com Peritos	13
5.1	Teste com Peritos em Tecnologia, Turismo, Saúde e Bem-estar	13
5.2	Teste de acessibilidade	20
5.3	Teste de Microcopy	21
6.	Considerações sobre Testes com Peritos	23
7.	Referências Bibliográficas	24
8.	Apêndice 1 - Grelha de validação heurística	25

# 1. Introdução

O presente *Deliverable* foi elaborado no âmbito do programa de investigação CeNTER e tem como objetivo apresentar os resultados das avaliações de uma plataforma digital piloto, de mediação, para a Inovação Territorial na Região Centro de Portugal. A partir da estruturação de um quadro de referências analítico, definidor de estratégias de estabelecimento/reforço de redes e de comunicação/mediação para impulsionar as dinâmicas de participação informal e envolvimento cívico das populações, o *WorkPackage 3 (WP3)* do CeNTER propõe o desenvolvimento de uma solução tecnológica digital, para a promoção da inovação de base territorial na Região Centro.

Este documento foi elaborado com o objetivo de reportar e disseminar as entregas públicas no âmbito do Projeto CeNTER, sendo o terceiro de quatro *deliverables* do WP3. Neste contexto, a primeira entrega pública apresentou um “Quadro de referência para estabelecimento de redes e estratégias de mediação”, e a segunda apresentou a proposta de “Implementação da plataforma piloto CeNTER”.

Por meio dos resultados das avaliações das diferentes versões desenvolvidas da referida plataforma piloto, a equipa do WP3 pretende validar a proposta elaborada, com vista a legitimar o modelo conceptual e tecnológico ao nível de estratégias de estabelecimento de redes e mediação para a inovação de base territorial, desenvolvida em fases anteriores, nomeadamente na Tarefa 3.1 do projeto.

Neste sentido, e tendo em conta que o *feedback* contínuo dos utilizadores em estádios iniciais do desenvolvimento de um produto tecnológico é crucial para detetar possíveis problemas que um sistema possa apresentar, foi dinamizada uma fase principal de testes que incluiu a avaliação do protótipo por peritos em contexto laboratorial. Tais testes abarcaram a apreciação do protótipo em diferentes fases de evolução e sob perspetivas variáveis, possibilitando uma avaliação mais completa. Os protótipos de baixa fidelidade foram anteriormente avaliados pela equipa interna do WP3, com vista a obter uma validação interna antes de passar para as etapas seguintes.

As alterações realizadas após a análise dos resultados das avaliações dos protótipos de média fidelidade da aplicação CeNTER deram origem a versões mais evoluídas da proposta digital do WP3. Este processo iterativo de desenvolvimento do protótipo foi alicerçado por um Design Centrado no Utilizador (DCU), metodologia de desenvolvimento que direcionou esta proposta para as necessidades de iniciativas e agentes da Região Centro de Portugal, identificadas após o estudo aprofundado de contextos de envolvimento cívico formais e informais relevantes, quer numa dimensão internacional, quer no território concreto da Região Centro (Tarefa 3.1).

Com a finalidade de adquirir os conhecimentos necessários para a realização dos testes da aplicação digital piloto CeNTER, foi realizado um levantamento teórico relativo à metodologia Design Centrado no Utilizador (DCU) e relativo ao conceito de *User Experience (UX)*. Além disso, foi feito um levantamento sobre os métodos existentes para validação de plataformas e aplicações *mobile*.

Desse modo, este documento apresenta-se da seguinte forma: o Capítulo 1 apresenta as considerações relevantes relativas ao DCU, enquanto os Capítulos 2 e 3 apresentam, respetivamente, informações relevantes sobre a UX e os testes conduzidos com Peritos pela equipa do WP3.

## 2. Design Centrado no Utilizador

Para responder ao último objetivo do projeto CeNTER, ou seja, à prototipagem e validação de uma plataforma piloto de inovação territorial para a Região Centro de Portugal, a equipa do WP3 desenvolveu um protótipo

de uma aplicação de telemóvel, ou *Mobile App* (m-app) utilizando como metodologia de desenvolvimento o Design Centrado nos Utilizadores.

O DCU tem como propósito definir o processo necessário ao desenvolvimento de produtos mais fáceis de utilizar e que, portanto, cumprem melhor os objetivos relativos à usabilidade (Fonseca et al., 2012). Inclui, deste modo, o envolvimento ativo dos utilizadores ao longo de todo o processo de desenvolvimento do produto ou serviço, por forma a evitar que os sistemas digitais falhem por falta de comunicação entre os desenvolvedores e os utilizadores (Still & Crane, 2017, p. 3).

Para Still & Crane (2017), os profissionais de design precisam de um conjunto de princípios norteadores no processo de desenvolvimento de um produto, de modo a que possam adaptá-lo conforme as necessidades de cada utilizador. Assim, existem mandamentos que garantem que se está a projetar um bom DCU como: *“Users are always right or you are never your user, don’t design for you, design for users.”* (p. 43).

Adicionalmente, Still & Crane (2017) apresentam 10 princípios fundamentais, necessários para aplicar na construção de experiências de DCU:

- Deve-se envolver os utilizadores desde o início do projeto;
- Deve-se envolver os utilizadores frequentemente ao longo do projeto;
- Deve-se projetar para uso em contexto real;
- Deve-se manter o produto, ou serviço, simples;
- Deve-se ser educado;
- Deve-se conhecer os utilizadores finais;
- Deve-se dar aos utilizadores controlo sobre o produto;
- Deve-se lembrar e projetar a emoção - as pessoas sentem tanto quanto pensam;
- Deve-se confiar, mas verificar - a triangulação é a chave;
- Deve-se investigar e descobrir antes de projetar e entregar, mas deve-se saber que essa descoberta nunca termina, mesmo após a entrega.

De acordo com estes autores, o cumprimento destes princípios torna possível desenvolver um produto ou serviço integralmente centrado no utilizador. No entanto, alguns tópicos relevantes são destacados. O princípio v: *“Ser educado”*, nas palavras de Still & Crane (2017) p. 54, quer dizer: *“In short, politeness builds user morale and makes the user feel in control. As a UCD designer then, you must be polite. Instructions are a perfect example.”*. Já o princípio ix: *“Deve-se confiar, mas verificar - a triangulação é a chave”* significa que, para entender os desejos e necessidades dos utilizadores, é preciso reunir o máximo de dados observáveis possíveis em todo o processo de design e, em seguida, fazer uma análise comparativa desses dados para apurar que semelhanças são encontradas. Para isso utilizam-se diferentes métodos de avaliação, que incluem a recolha de dados qualitativos e quantitativos. Assim, é possível detetar o maior número de erros e realizar as alterações necessárias para melhorar o protótipo.

Em termos de processo e metodologia de desenvolvimento do DCU, o princípio essencial está resumido na palavra *“iteração”*. *“Design Thinking its iterative in nature”* (Still & Crane, 2017, p. 4). Neste contexto, a Figura 1 mostra o processo iterativo do DCU apresentado por Don Norman (2013), que envolve, primeiramente, a observação do utilizador para tentar compreender seus interesses, motivos e verdadeiras necessidades. Assim, pretende-se conhecer quem irá utilizar o sistema interativo e que funcionalidades poderão dar suporte melhor às suas tarefas. Posteriormente, há a idealização de soluções de design através do desenvolvimento de modelos conceptuais e criação de sketches em papel, e de mockups, para demonstração e validação inicial do projeto delineado. A fase seguinte inclui a evolução do modelo a partir do desenvolvimento de um protótipo, que será testado na etapa seguinte, onde se definem as modificações que forem necessárias para adequação, ainda no protótipo. Ressalta-se que o sistema iterativo inclui repetidos testes e avaliações da solução, até se alcançar a satisfação das necessidades e requisitos dos utilizadores finais (Norman, 2013).

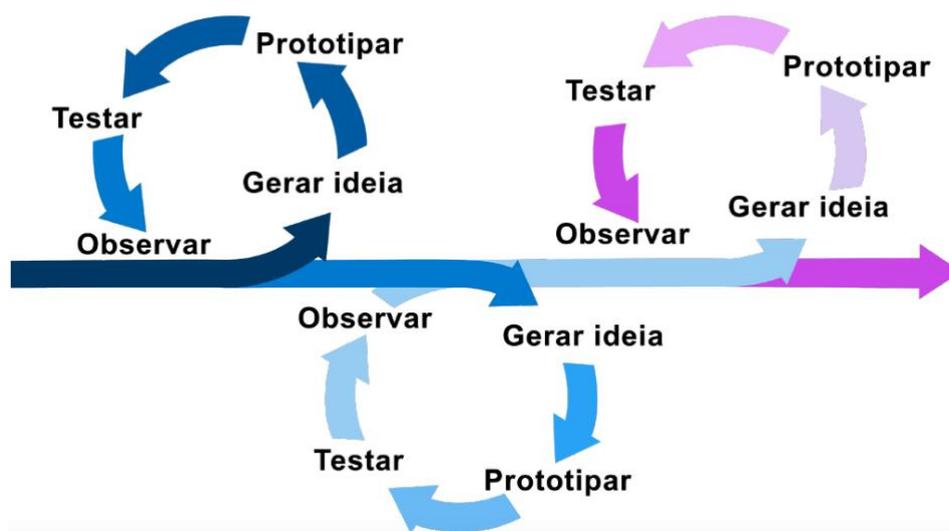


Figura 1 – Processo iterativo do Design Centrado no Utilizador (Norman, 2013).

Por ser um processo que pretende ir ao encontro das necessidades e exigências dos utilizadores, o DCU é frequentemente associado ao sucesso de um produto e/ou, uma ótima UX, uma vez que bons projetos de design são frequentemente acompanhados por altos padrões de desempenho e boas métricas de usabilidade (Knight, 2019). A próxima secção apresenta os conceitos mais relevantes sobre a UX, bem como as características da mesma que contribuem para que um produto obtenha sucesso, sendo por isso pertinente no âmbito do desenvolvimento do protótipo CeNTER.

### 3. Avaliação da User Experience

É possível encontrar diversas definições para *User Experience* (ou Experiência do Utilizador) na literatura, não sendo o seu significado ainda estável e consolidado. O termo foi criado por Don Norman e abrange toda a experiência que um indivíduo tem, através de todo e qualquer meio, em torno de um produto ou serviço (Knight, 2019). O termo *User Experience* refere-se à forma como o utilizador final se sente sobre os produtos criados (idem, p. 2). A experiência é, assim, um constructo formado na própria mente, para além de uma infinidade de outros fatores (idem, p. 2). De acordo com Knight (2019), a UX é uma questão completamente subjetiva. Já Bernhaupt e Pirker (2013) afirmam que o conceito de UX está relacionado com emoções positivas e resultados emocionais, como alegria, diversão e orgulho.

Para Knight (2019), a criação de uma experiência não é apenas sobre a forma como o produto é criado. Não se trata de como foi projetado, quais as estruturas que foram implementadas ou se é utilizada uma tecnologia de ponta. É sobre como o produto ajuda as pessoas a realizarem as suas tarefas, alcançar os seus objetivos e como se sentem quando usam e se envolvem com o produto. No caso das soluções digitais, por exemplo, as intenções são transformadas em produtos, que serão usados por pessoas reais. Neste sentido, a abordagem Design Centrada no Utilizador apresenta-se como uma opção:

This user-centered mindset can be a commonly neglected aspect of digital product design and is one that needs to receive far more prominence. By providing developers with the means to gain a comprehensive understanding of your user's needs, and what good user experience means to your users, we can create more successful digital products with a solid foundation of knowing their needs and goals. (Knight, 2019, p. 2).

Knight (2019) apresenta seis categorias principais da UX, sistematizadas num conjunto de especialidades que se concentram nas áreas de criação de produtos digitais:

- *User Research* – compreender os utilizadores e as suas necessidades e exigências;
- *Content strategy* – garantir que o sistema/produto apresenta conteúdos significativos e atraentes;
- *Information Architecture* – as informações devem estar organizadas de modo lógico, sustentável e claro;
- *Interaction design* – providencia uma interação fluida e atraente com a interface, através de pistas visuais, respostas às ações dos utilizadores através de *feedbacks* e antecipação de erros;
- *Visual design* – refina a tipografia, cores, imagens e outros elementos visuais, para ajudar a transmitir o conteúdo ou a função do produto;
- *Usability* - a usabilidade refere-se à eficiência, eficácia e satisfação do utilizador ao interagir com um produto digital. Diferentes técnicas podem ser implementadas para determinar quão intuitivo é o design, descobrir a frequência de erros ao executar tarefas, revelar como é fácil aprender e usar e, até, medir o *feedback* subjetivo em termos de satisfação do utilizador.

Uma outra categorização da UX foi realizada por Hassenzahl e Tractinsky (2003), na qual o conceito de UX é organizado em dimensões, como mostrado abaixo:

- Qualidades não-instrumentais: A este nível, importa considerar dimensões como a estética das aplicações, que, pela perceção da experiência, se pode traduzir em sensações de surpresa, diversão e, eventualmente, intimidade, satisfazendo necessidades humanas que ultrapassam o conceito instrumental e funcional da usabilidade. Outros conceitos, que se enquadram nesta categoria, podem ser resumidos como aspetos hedónicos da utilização da tecnologia, como estimulação e identificação;
- Reações emocionais: nesta vertente da UX, a tendência atual centra-se em torno das emoções positivas e dos resultados emocionais, como nível de prazer, atração e controlo sobre a aplicação;
- Prática temporal: A terceira vertente relaciona-se com as influências temporais e situacionais, procurando verificar quão única, complexa, temporária ou situada é a experiência ao nível dos estados internos do utilizador (por exemplo, humor, expectativas, objetivos alcançados).

Para além das dimensões acima apresentadas, os componentes de usabilidade são conhecidos, também, como qualidades instrumentais que englobam aspetos como eficiência, eficácia, facilidade de aprendizagem e de recordação de um sistema, bem como da sua utilidade (Hassenzahl & Tractinsky, 2003). De acordo com Fonseca et al. (2012), quando a usabilidade de um sistema é avaliada, procura-se identificar qual a facilidade de aprender, memorizar e usar o sistema e, perante o mesmo, qual é a sua eficiência, eficácia e segurança. Assim, idealmente, a avaliação de usabilidade deve estar presente em todas as etapas do processo de design e desenvolvimento, e deve ser iterativa de forma a permitir a melhoria contínua dos resultados (Martins et al., 2015).

A partir das qualidades instrumentais e não-instrumentais é possível elencar diversos instrumentos para aferição de características da UX nas diferentes áreas que a mesma abrange. Para além disso, de acordo com Knight (2019), a produção de um produto deve considerar a UX em vários passos, ao longo do desenvolvimento do projeto de design.

Desse modo, a avaliação no âmbito do Projeto CeNTER foi alicerçada por uma abordagem iterativa, sendo que o primeiro momento incluiu a avaliação por parte de peritos, seguido de uma série de avaliações personalizadas aos diferentes agentes da comunidade, i.e., Iniciativas de Base Comunitária, Cidadãos, Entidades Públicas e Privadas e Redes. Da mesma forma, a avaliação da usabilidade do protótipo foi realizada

em todas as etapas de testes conduzidos. As próximas seções apresentam as aferições do protótipo da *m-app* realizados no âmbito do projeto CeNTER.

## 4. Teste com Peritos

O processo da avaliação da *User Experience* de um protótipo preliminar pode incluir a avaliação por especialistas, uma etapa crucial na identificação de problemas relacionados com a usabilidade do produto e na recolha de informações relevantes sobre a *User Interface* (UI). Para além disso, identifica-se como importante para avaliar os requisitos dos utilizadores, a viabilidade das funcionalidades do sistema e ainda ao nível da validação de conceito. Neste sentido, os peritos locais são fontes importantes de informação, pois apresentam conhecimento e experiência sobre o cenário final de utilização do produto, auxiliando na compreensão sobre as necessidades do utilizador final e restrições locais ao desenvolvimento do sistema (Still & Crane, 2017).

Para que a realização dos testes seja feita de modo eficaz é necessário que os avaliadores sejam escolhidos de acordo com a sua área de especialização, considerando a área profissional e a experiência com sistemas similares. Do mesmo modo, é importante utilizar métodos que possibilitem uma visão global da usabilidade da interface e que permitam aceder às principais funcionalidades disponíveis no protótipo. Para realizar a avaliação do protótipo do *mobile-app* CeNTER, dois painéis de especialistas foram organizados, de acordo com a experiência dos avaliadores.

### 4.1 Objetivos dos Testes

Com vista a orientar as fases seguintes de desenvolvimento do protótipo, os testes com especialistas visaram essencialmente avaliar as principais funções das interfaces do protótipo. Assim, os objetivos da avaliação com especialistas foram:

- Validar o conceito do protótipo CeNTER;
- Verificar a aceitação das interações dentro do protótipo (modo de interação *drag and drop* e *swipe*, por exemplo);
- Avaliar o *look & feel* das interfaces;
- Avaliar a usabilidade do protótipo;
- Recolher sugestões de melhorias;
- Identificar erros de prototipagem;
- Recolher outros dados importantes acerca de aspectos relacionados à UX, nomeadamente informações sobre as reações emocionais e aspectos hedónicos do uso do protótipo;
- Verificar se o protótipo apresenta os requisitos de acessibilidade;
- Identificar se o conteúdo textual utilizado no protótipo do aplicativo CeNTER é adequado e correto, levando em consideração o contexto do projeto;
- Apoiar a validação lexical e terminológica dos conteúdos utilizados na aplicação.

## 4.2 Metodologia

### 4.2.1 Teste com Peritos em Tecnologia e Turismo, Saúde e Bem-estar

O processo de avaliação da UX no âmbito do CeNTER baseia-se num modelo espiral de design iterativo, incluindo testes iniciais dos modelos de baixa fidelidade com os membros WP3, avaliação do protótipo de média fidelidade por especialistas e testes laboratoriais com utilizadores finais. Para atingir os objetivos do teste com Peritos e maximizar a identificação de problemas de usabilidade do protótipo, durante o processo de avaliação, foram utilizadas duas técnicas:

1. Técnica de inspeção baseada em avaliação heurística: apoiada em lista e itens de verificação, que permite aos avaliadores investigarem e identificarem problemas de usabilidade nas interfaces do sistema, não envolvendo diretamente o utilizador final;
2. Técnica de observação, na qual uma visão exploratória do protótipo usando um protocolo *Think-aloud* foi incentivado. Neste caso, as interações e comentários dos avaliadores durante a realização do teste foram registados pela equipa.

O processo de avaliação heurística ocorreu em três etapas, conforme proposta de Nielsen (1994):

- Fase de preparação na qual foram definidos e organizados o protótipo de avaliação e a lista de heurísticas a serem utilizadas;
- Fase de avaliação que consistiu na recolha de dados de avaliação com cada perito, individualmente. Os avaliadores testaram o protótipo, identificando as diretrizes ausentes e o grau de gravidade do problema. As avaliações foram feitas individualmente e em igualdade de condições. Os dados da avaliação foram recolhidos usando a mesma lista de verificação heurística e registo de observação;
- Fase de consolidação dos resultados e relatório dos resultados que decorreu no final dos testes, tendo os dados recolhidos dos dois painéis de avaliadores sido analisados e comparados.

Foi desenvolvida uma lista de verificação (Apêndice 1), suportada no conjunto de heurísticas mais relevantes de serem testadas ao nível do protótipo CeNTER, ou seja, foram escolhidos apenas itens de verificação que o *software* “Principle”, usado para a prototipagem da *m-app*, permitisse testar. Neste sentido, a lista incluiu dez heurísticas de Nielsen (1994) e três heurísticas da escala MATCH, orientadas para sistemas *mobile* (Salazar et al., 2012). Assim, um total de 44 itens estavam presentes na lista de verificação para avaliar a usabilidade do protótipo *m-app* CeNTER.

Na fase de avaliação, o processo começou com a apresentação do Projeto CeNTER e, posteriormente, os avaliadores foram orientados e incentivados a explorar livremente o protótipo da aplicação. A avaliação consistiu em duas fases: a primeira, na qual os especialistas exploraram livremente o protótipo e comentaram ou responderam às suas dúvidas sobre o “CeNTER”, e a segunda, que incluiu o preenchimento da tabela de heurísticas orientadas ao design de interfaces de aplicações *mobile*. Algumas modificações no protótipo foram feitas após a análise dos resultados obtidos pelo primeiro painel de avaliação. A versão testada pelo segundo painel foi, deste modo, diferente da primeira e incluiu as primeiras melhorias feitas pelo WP3 ao nível da interface e funcionalidades.

### 4.2.2 Teste com Perito em Acessibilidade

A avaliação da acessibilidade foi realizada por meio de avaliação especializada e teste de ferramentas digitais. O teste com o especialista em acessibilidade teve quatro etapas. i) Preparação: Diversas ferramentas digitais gratuitas também foram exploradas, que permitem gravar a tela do celular do especialista (por exemplo, TeamViewer, gravador de tela do iPhone). Depois disso, a ferramenta ApowerMirror foi selecionada de acordo com sua compatibilidade com o iPhone e o software Principle, permitindo-nos capturar a experiência do usuário em todas as telas. O ApowerMirror nos permite compartilhar a tela do celular no notebook. A

plataforma Zoom foi selecionada para videoconferência, para compartilhamento da tela do especialista e registo do processo de avaliação; ii) primeiro contato com o especialista: foi enviado um e-mail convidando o especialista a participar dos testes. Após a aceitação do convite, os arquivos com o protótipo e a solicitação de consentimento foram enviados ao avaliador. iii) sessão de pré-teste (remoto): Configure toda a tecnologia necessária para o teste (instale o aplicativo Principle e ApowerMirror); iv) teste de acessibilidade (remoto): O teste foi de forma moderada, com a presença de um especialista em acessibilidade e pesquisadores da equipe do CeNTER.

O teste foi em 28/05/2020 e teve duração de 3 horas e 30 minutos. A sessão começou com o especialista em acessibilidade sendo convidado a explorar a tela aberta e o menu a seguir. A especialista foi solicitada a comentar o que estava procurando ou lendo, o que gostou e não gostou no aplicativo e o que o tornou fácil ou difícil de executar. A solução técnica foi a especialista compartilhar a tela do seu telefone em seu computador, através de wi-fi e software ApowerMirror. Depois disso, o especialista usou o Zoom para compartilhar a tela do computador com os membros da equipe do CeNTER. Essa solução técnica permitiu acompanhar todas as etapas de navegação e experiência do usuário no protótipo.

Os resultados foram transcritos e analisados em grade, definindo os níveis de prioridade e complexidade dos problemas detetados. Todas as demandas foram priorizadas para posterior desenvolvimento. A partir daí, os problemas foram revistos, corrigindo os problemas identificados e implementando as respetivas melhorias.

### **4.2.3 Teste de Microcopy**

O Teste de Conteúdo de Microcópia visa identificar se o conteúdo textual utilizado no protótipo do aplicativo CeNTER é adequado e correto, levando em consideração o contexto do projeto. Este processo teve como objetivo apoiar a validação lexical e terminológica dos conteúdos utilizados na aplicação. Todos os 215 elementos textuais apresentados nas interfaces foram avaliados quanto à adequação de cada conteúdo textual, em escala Likert, e se o conteúdo foi utilizado corretamente ou não. Além disso, o especialista foi convidado a apresentar sugestões de alteração dos elementos textuais que, em sua opinião, deveriam ser modificados. Os dados foram posteriormente analisados pela equipe do CeNTER.

Para a avaliação foi construída uma grelha com todos os conteúdos lexicais apresentados nas interfaces do protótipo CeNTER. Esta grade de conteúdo da microcópia foi enviada por e-mail ao especialista juntamente com um documento pdf com as imagens das interfaces do protótipo. Portanto, o teste de Microcopy não foi moderado por nenhum membro da equipa. Porém, é importante destacar que todas as questões foram respondidas por e-mail e que a equipa estava sempre disponível para esclarecer qualquer dúvida que pudesse surgir.

## **4.3 Perfil dos Avaliadores**

Foram organizados dois painéis de especialistas, de acordo com a experiência dos avaliadores. O primeiro painel foi composto por cinco especialistas na área das tecnologias digitais, que possuem conhecimento e experiência no desenvolvimento de interfaces. Estes peritos são membros integrados do DigiMedia, do Departamento de Comunicação e Arte da Universidade de Aveiro. O segundo painel foi composto por cinco especialistas nas áreas de Turismo, Saúde e Bem-Estar, que conhecem o domínio de aplicação da tecnologia que está a ser desenvolvida e estão envolvidos em diferentes projetos comunitários na Região Centro de Portugal. As avaliações foram realizadas, individualmente, com cada perito. O segundo painel de peritos avaliou o protótipo 11 dias após a finalização dos testes pelo primeiro painel, ambos realizados com base no mesmo protocolo de avaliação.

Relativamente ao teste com perito em acessibilidade, o avaliador foi uma estudante de doutoramento da Universidade de Aveiro, enquanto o teste de microcopy foi realizado com um Professor de Português e Linguística da Universidade de Londres e da Universidade de Aveiro.

## 4.4 Ambiente dos Testes

Os testes foram realizados nas instalações da Universidade de Aveiro, com o primeiro painel a realizar os testes entre 28 e 31 de outubro de 2019 e o segundo painel entre 11 e 29 de novembro de 2019. No total, foram avaliados 78 ecrãs do protótipo CeNTER desenvolvido. Em ambos os casos, a avaliação foi realizada através de um dispositivo *mobile* iPhone 8, onde os ecrãs eram transferidos do software “Principle”, a partir de um iMac. As avaliações não tiveram uma duração pré-definida, mas duraram, em média, cerca de 1h30.

No que se refere aos testes de acessibilidade e microcopy, ambos ocorreram remotamente. No entanto, enquanto o teste de acessibilidade incluiu o uso de ferramentas e digitais para que o teste fosse monitorado pela equipa da WP3 do CeNTER, o teste de Microcopy não foi monitorado, de modo que as grelhas de avaliação desenvolvidas fossem enviadas por e-mail para o avaliador.

## 4.5 Versões do Protótipo

Após a tarefa inicial de levantamento de requisitos e identificação de diretrizes para a solução digital de inovação territorial da Região Centro de Portugal, o desenvolvimento do protótipo CeNTER incluiu, num primeiro momento, diversos *brainstorms* entre a equipa da WP3 para a validação interna dos primeiros esboços elaborados. Após os primeiros desenhos em papel, foram criados *mockups* de baixa fidelidade, aprovados após reuniões da equipa.

Tendo em conta os requisitos e funcionalidades definidos como importantes na solução digital proposta, o protótipo de média fidelidade foi elaborado tendo como suporte o “Principle”, um *software* de design de *User Interfaces* interativas e animações. O enriquecimento das funcionalidades e da interface, ao longo do tempo, contribuiu para que o protótipo adquirisse robustez, de modo a que deu origem a versões anteriores aos primeiros testes com peritos.

Estruturalmente, o conteúdo do protótipo que foi para testes com peritos foi organizado de acordo com as categorias definidas pela equipa como sendo essenciais na solução digital no âmbito do CeNTER: Iniciativas; Atividades; Entidades; Voluntários; Recursos e Destaques. Tais categorias foram organizadas em linhas horizontais no ecrã e apresentam *clusters* de conteúdos que variam de acordo com as preferências definidas nas configurações da *app*. Ao tocar nas *tags*, ocorre uma abertura vertical no ecrã, onde será possível visualizar as sugestões recomendadas através de *cards* ordenados em carrossel. Para além disso, cada categoria apresenta opções de visualização no mapa, adicionar um novo elemento e busca específica no âmbito de cada *tag*. A Figura 2 apresenta o ecrã definido como *Homescreen*, com os conteúdos das diferentes categorias.

O cabeçalho do protótipo apresenta, para além do nome do projeto e da logomarca do CeNTER, funcionalidades de agenda, pesquisa geral em toda a plataforma, e acesso às configurações. O *footer* é composto por outras opções de funcionalidades, tais como aceder ao perfil a às ideias, visitar os itens guardados, visualização de notificações e exploração geral no mapa.



Figura 2 - Homescreen da versão do protótipo de média fidelidade “CeENTER” avaliada por peritos.

A versão a seguir ao primeiro painel de testes incluiu ligeiras modificações em comparação com a versão anterior. Um detalhe maior das funcionalidades e características de design das respetivas versões dos protótipos pode ser visualizado no *Deliverable* 2.3 elaborado pela equipa da WP3.

### 4.5.1 Versão para o Teste de Acessibilidade e Microcopy do Protótipo CeENTER

Após a realização dos dois painéis de testes com peritos, algumas modificações foram feitas relativamente aos conteúdos e ao *look & feel* das interfaces. As principais modificações / alterações no protótipo foram realizadas com vista a dar respostas a uma série de sugestões e erros de prototipagem no âmbito dos testes com peritos. Algumas dessas modificações são apresentadas a seguir e podem ser observadas na Figura 3:

- foi desenvolvida uma animação no *software* “Principle”, com o objetivo de informar sobre as principais funcionalidades da plataforma e clarificar as diferenças entre as categorias essenciais no âmbito do CeENTER: Iniciativas; Atividades; Entidades; Voluntários; Recursos e Destaques;
  - esta modificação foi uma resposta face às dúvidas relativas ao significado de cada categoria, principalmente a diferença entre “Atividade” e “Iniciativa”. A categoria “Atividade” foi substituída por “Evento”, ao passo que se mantiveram as outras nomenclaturas. Neste sentido, a animação apresenta-se como uma importante melhoria.
- o tutorial foi modificado no sentido de aumentar e ressaltar as letras utilizadas para o texto informativo na parte mediana do ecrã. O espaço entre os textos no ecrã (textos na parte inferior e mediana do ecrã) foi alargado para aumentar a nitidez dos conteúdos. Ainda sobre o tutorial, as imagens foram padronizadas em termos de tamanho e lugar nos ecrãs;
  - estas modificações foram realizadas tendo em conta alguns comentários, ao longo dos testes, acerca do quão inteligível eram, tanto as imagens, quanto os textos descritivos na parte inferior dos ecrãs. Assim, verificou-se a necessidade de desenvolver outras interfaces, com o corpo textual e imagético mais nítido, compreensível e padronizado;
- O *Homescreen* foi reconfigurado de modo que a funcionalidade relativa ao mapa esteja sempre visível na parte superior às categorias supracitadas.
  - esta re-configuração foi sugerida por um dos peritos que compôs o painel em tecnologias digitais. Este perito identificou o “mapa” como uma função que valoriza a plataforma e que

deveria estar sempre à vista dos utilizadores, de modo que seja fácil obter informações acerca dos eventos que ocorrem em locais próximos e em tempo real.

- Ainda no *Homescreen*, as funcionalidades que são apresentadas no cabeçalho, à direita do ecrã, foram reordenadas e substituem a configuração anterior (“Definições; Agenda e Lupa”) do seguinte modo: “Agenda, Lupa e Definições”. No entanto, as categorias, do mesmo modo que a versão anterior, foram organizadas em linhas horizontais no ecrã e apresentam clusters de conteúdos que variam de acordo com as preferências definidas nas configurações da *app*. Ao tocar nas *tags*, ocorre uma abertura vertical no ecrã, onde será possível visualizar as sugestões recomendadas através de *cards* ordenados em carrossel. Para além disso, cada categoria apresenta opções de visualização no mapa, adicionar um novo elemento e busca específica no âmbito de cada *tag*.

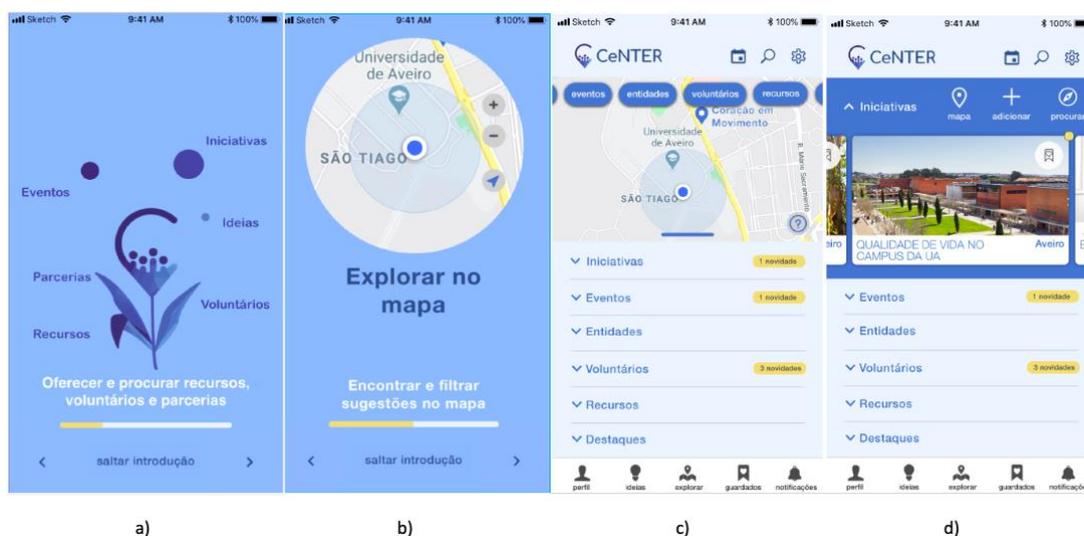


Figura 3 - Interfaces da versão do protótipo de média fidelidade “CeNER” após as avaliações realizadas pelos dois painéis de peritos. a) animação informativa mostrada no primeiro uso da plataforma. b) exemplo de interface do tutorial com padronização das imagens, em círculos, e reconfiguração do tamanho e locais das letras no ecrã. c) *Homescreen*, com a funcionalidade mapa na parte superior no ecrã e categorias apresentadas em linhas horizontais em baixo. d) *Homescreen* com categorias organizadas em linhas horizontais no ecrã, com clusters de conteúdos apresentados em *cards* ordenados em carrossel que se abrem verticalmente.

## 5. Resultados dos Testes com Peritos

### 5.1 Teste com Peritos em Tecnologia, Turismo, Saúde e Bem-estar

Considerando os principais objetivos dos testes realizados, as recomendações dos especialistas comprovaram que uma aplicação de promoção da inovação de base territorial é uma proposta sustentável e oportuna, embora ainda existam algumas questões que serão abordadas posteriormente neste documento. Os resultados são apresentados de acordo com a técnica de avaliação utilizada: observação, relatos e técnicas heurísticas. O relatório das interações e comentários dos especialistas por meio do protocolo *Think-aloud* durante os testes tornou possível extrair resultados qualitativos e quantitativos, enquanto a avaliação heurística forneceu dados quantitativos valiosos. Para definir as principais melhorias a serem realizadas no protótipo, os dados registados foram separados em duas tabelas<sup>1</sup>, de acordo com o painel correspondente

<sup>1</sup> Tabela com *inputs* do painel 1 – <https://docs.google.com/spreadsheets/d/12Vh0ANB-KGXnNnEuY3gKEqXgfvvTpdAACSzGpmEH0s/edit?usp=sharing>; Tabela com *inputs* do painel 2 - <https://docs.google.com/spreadsheets/d/17682ldfMR-8x1GNuZdMhXDKFOEUAgfR1tgaSizDMzsl/edit?usp=sharing>

do teste. Posteriormente, cada *input* foi medido de acordo com o nível de prioridade de modificação e o nível de esforço necessário para a realização da mesma. Os resultados são descritos seguidamente.

### 5.1.1 Resultados da Técnica de Observação

Após a realização dos testes com os dois painéis de avaliadores, os dados registados foram reunidos numa tabela Excel e os *inputs* repetidos foram descartados. De seguida, as informações foram classificadas de acordo com as suas semelhanças, ou seja, de acordo com os ecrãs correspondentes, com o especialista que fez o comentário, com a prioridade de cada item a ser modificado no protótipo e com o nível de esforço da respetiva modificação. Para além disso, os dados foram analisados de acordo com a natureza dos *inputs*, tendo sido classificados como “erros de prototipagem” ou “sugestões de melhoria”. Esta distribuição dos *inputs* forneceu informações valiosas sobre o *feedback* dos especialistas, permitindo alcançar, juntamente com a Avaliação Heurística, os objetivos do teste do protótipo. A análise final das grelhas deu origem às modificações mais relevantes a serem feitas para prosseguir com a prototipagem da aplicação CeNTER.

Cento e vinte e um (121) *inputs* foram relatados durante a exploração pelos especialistas em tecnologias digitais, 83 dos quais foram considerados pela equipa como sugestões de melhoria na plataforma, 30 como erros de prototipagem, e 8 foram interpretados como sugestões de melhoria e erros de prototipagem. Os erros de prototipagem correspondem a inconsistências no uso do protótipo e da interface, como ícones diferentes que representam a mesma função ou a falta de um ícone que deve estar presente por *default* (por exemplo, a ausência do ícone de retorno em alguns ecrãs, com consequente impossibilidade de executar esta ação através do modo considerado padrão no protótipo). Um exemplo de um erro de prototipagem é: *“Padronizar os ícones de ‘guardar’, no lado direito do ecrã ideias, com o restante da aplicação.”*. (Painel 1 - Especialista 1). Um exemplo de sugestões de melhoria é: *“Em ‘Ideias’, adicionar informações de ano e não apenas dia e hora”*. (Painel 2 - Especialista 2).

Em relação aos *inputs* recolhidos na avaliação dos especialistas nas áreas de Turismo, Saúde e Bem-Estar, foram obtidos 53 *inputs* durante a exploração do protótipo, 41 considerados pela equipa do CeNTER como sugestões de melhorias na plataforma, 11 como erros de prototipagem e 1 como sugestão de melhoria e erro de prototipagem. Nesse caso, um exemplo de erro de prototipagem é: *“Uma pessoa pode-se perder ao alternar entre guias, sem saber como voltar ao abrir uma guia e como minimizar outra.”*. (Painel 2 - Especialista 3). Um exemplo de sugestões de melhoria é: *“Altere ‘desporto’ por ‘atividade física e desporto’. O desporto é interpretado como ação competitiva”*. (Painel 2 - Especialista 4). A Tabela 1 apresenta o número de *inputs* registados pela equipa do CeNTER nas duas avaliações.

Verifica-se uma diferença considerável no número de *inputs* registados entre os dois painéis. Provavelmente isso ocorreu devido à diferença no carácter de especialização dos peritos que compuseram cada painel. Tendo em conta que o painel 1 era formado por especialistas em tecnologias digitais, para além de considerações pessoais sobre o conceito, funcionalidade, inovação e outros princípios, os mesmos observaram atentamente os aspetos técnicos do protótipo. Por outro lado, os especialistas do painel 2 comentaram principalmente o conceito do protótipo CeNTER, funcionalidade e eficácia para o utilizador final, enquanto os fatores técnicos não foram o alvo principal da sua apreciação. Um exemplo de situação que clarifica essa premissa é o facto de que apenas os especialistas da área das tecnologias digitais foram capazes de explorar a interação *drag and drop* sem nenhuma pista da equipa do CeNTER.

Tabela 1. Número de inputs dos Peritos dos Painéis 1 e 2 (P1 e P2)

<b>Inputs</b>	<b>P1</b>	<b>P2</b>	<b>Total</b>
Total de <i>inputs</i>	121	53	174
Sugestão de melhorias	83	41	124
Erros de prototipagem	30	11	41
Erros e sugestões	8	1	9

Os dados extraídos nas avaliações foram também analisados de acordo com a prioridade de cada item ser modificado no protótipo. Portanto, no primeiro painel de testes, 21 *inputs* foram considerados como prioridade 1 pela equipa, já que foram considerados problemas que afetariam a usabilidade do protótipo, influenciando a qualidade da experiência do utilizador. Quarenta e dois foram considerados prioridade 2, ou problemas que alteram a usabilidade do sistema, mas não prejudicam a qualidade da experiência do utilizador e 58 foram considerados prioridade 3, pois não afetam a qualidade da experiência do utilizador. Das contribuições de especialistas do segundo painel, 14 foram consideradas prioridade 1 pela equipa do WP3; 12 foram considerados prioridade 2; e 27 foram considerados prioridade 3 (Tabela 2).

Tabela 2. Prioridade dos inputs

<b>Prioridade dos <i>inputs</i></b>	<b>Nº de <i>inputs</i> P1</b>	<b>Nº de <i>inputs</i> P2</b>
Prioridade 1	22	14
Prioridade 2	41	12
Prioridade 3	58	27
Total	121	53

O nível de esforço para a melhoria do protótipo foi também avaliado. Assim, no primeiro painel de testes, 57 *inputs* apresentavam um baixo nível de esforço de implementação, pois requeriam um curto período de tempo para solucionar o erro/problema. Quarenta e dois dados foram considerados como um nível médio de esforço, pois exigiam algum tempo para a resolução do erro/problema, e 22 foram considerados como um nível de esforço alto, ou seja, exigiam muito tempo para a resolução do erro/problema (Tabela 3).

Tabela 3 – Nível de esforço para a melhoria no protótipo

<b>Nível de esforço</b>	<b>Nº de <i>inputs</i> P1</b>	<b>Nº de <i>inputs</i> P2</b>
Baixo nível de esforço	57	9
Médio nível de esforço	42	13
Alto nível de esforço	22	31
Total	121	53

Os *inputs* foram, ainda, divididos de acordo com cada ecrã correspondente. A Tabela 4 mostra o número de *inputs* relacionados com os principais ecrãs testados.

Tabela 4 – *Inputs* de acordo com a interface do protótipo

Interfaces	Nº. de <i>inputs</i> P1	Nº. de <i>inputs</i> P2
Tutorial	11	2
Ecrã inicial	21	19
Perfil	10	3
Registo de iniciativa /atividade	16	4
Ideias	16	2
Mapas	7	3
Agenda	8	1
Favoritos	5	1
Notificações	6	2
Ecrã detalhes	16	9
Outros	5	7
Total	121	53

Por fim, os *inputs* dos dois painéis foram divididos em erros de prototipagem e sugestões de melhoria, de modo que são, ainda, identificados de acordo com as categorias de afinidade<sup>2</sup>. A Tabela 5 apresenta as sugestões de melhoria e os erros de prototipagem, de acordo com as respetivas categorias.

Como dito anteriormente, com a intenção de determinar as principais modificações a serem feitas para a melhoria do protótipo, todas as informações acima foram separadas em duas tabelas diferentes, de acordo com o painel de teste correspondente. Para escolher as alterações a serem feitas no protótipo pela equipa do CeNTER, os principais fatores levados em consideração foram o nível de prioridade de cada item registado e o nível de esforço para a melhoria no protótipo. Assim, os *inputs* dos especialistas foram avaliados no nível de prioridade e nível de esforço da respetiva modificação descrita. Os itens com alto nível de prioridade e baixo nível de esforço foram considerados como prioritários para ajuste. Assim, embora todos os resultados de cada painel sejam apresentados em conjunto neste documento, tal é feito apenas para fins de comparação, pois foram analisados separadamente, sendo que cada painel deu origem a modificações que refletem os *inputs* de cada conjunto de especialistas.

Após a análise das tabelas, o primeiro painel de testes deu origem a 14 itens a serem modificados no protótipo, todos exigindo baixo nível de esforço. Seis itens apresentam alto nível de prioridade (prioridade 1), sete apresentam nível médio de prioridade (prioridade 2) e um apresentou baixo nível de prioridade (prioridade 3). Na tabela do segundo painel de testes, 20 itens foram selecionados para serem modificados no protótipo da aplicação. Treze itens apresentam alto nível de prioridade (prioridade 1), seis apresentam nível médio (prioridade 2), e um apresenta baixo nível de prioridade (prioridade 3). Quanto ao esforço

<sup>2</sup> Link para a tabela completa:

[https://docs.google.com/spreadsheets/d/112WYyPnBllm7m8vj58ZNtgH4N2dvPNEcKAttLc5\\_zho/edit?usp=sharing](https://docs.google.com/spreadsheets/d/112WYyPnBllm7m8vj58ZNtgH4N2dvPNEcKAttLc5_zho/edit?usp=sharing)

necessário, 7 requeriam baixo nível de esforço, enquanto 7 e 6 exigiam, respetivamente, um nível médio e alto de esforço.

*Tabela 5 - Tabela de erros de prototipagem e sugestões de melhorias por categorias*

<b>Categorias</b>	<b>Nº de erros de prototipagem</b>	<b>Categorias</b>	<b>Nº de sugestões de melhorias</b>
Problemas na Interação <i>drag and drop</i>	3	Hierarquia e organização das interfaces	7
Inconsistências	16	Sugestões na Interação <i>drag and drop</i>	6
Separadores	3	Inconsistências	9
Nomenclatura / compreensão	6	Sugestões de nomenclaturas / trocas de designações	23
<i>Toogles</i>	2	Reconfiguração dos ecrãs	13
Realocação de funcionalidades	3	Otimização da visibilidade da informação	23
Hierarquia e organização das informações na interface	5	Sugestões de funcionalidades não encontradas	20
Apoio ao utilizador	2	<i>Feedback</i> da app para o utilizador	10
Design minimalista	4	Reconfiguração de cores	4
Reconfigurações	6	Botões e ícones	5
-	-	Sugestões ao CeNTER	5
-	-	Minimização de informações	8
<b>Total</b>	<b>50</b>		<b>133</b>

Exemplos de melhorias selecionadas no primeiro painel de testes são: "*Deve-se poder voltar para um ecrã anterior no tutorial*". (Painel 1 - Especialista 1), "*A função de pesquisa no ecrã 'ideias' deve ser consistente com o restante da aplicação*". (Painel 1 - Especialista 3 e 4), "*A ordem dos ícones de 'pesquisa' e 'configurações' no ecrã de 'notificações' deve ser mudado, levando em consideração o que você vê no ecrã inicial ou no mapa*". (Painel 1 - Especialista 5) e "*Padronização dos ícones salvos*", no lado direito do ecrã ideias, com o restante da aplicação". (Painel 1 - Especialista 5).

Exemplos das melhorias selecionadas no segundo painel de testes são: "*Deveria ser melhor hierarquizado/explicado para entender o conteúdo presente na plataforma (iniciativas, atividades, voluntários, recursos, etc.)*" (Painel 2 - Especialista 4), "*Sugere-se encontrar outra nomenclatura para*

*iniciativas, pois elas são confundidas com as atividades” (Painel 2 - Especialista 1, 2, 3 e 4), “O gesto de arrastar e soltar é muito rápido. Parece não ter controle sobre os elementos de interação” (Painel 2 - Especialista 1 e 4) e “Na tela ‘partilhar’: qual é a diferença entre os ícones SMS e Telefone?” (Painel 2 - Especialista 1, 2, 3 e 4).*

Embora haja um contraste considerável no número de *inputs* relatados entre os dois painéis, com o painel 1 a apresentar uma quantidade maior, alguns *inputs* merecem ser destacados, pois foram apontadas pelos dois painéis. Globalmente foi consensual que o sistema de interação *drag and drop* é um desafio do ponto de vista da usabilidade. Somente os especialistas da área das tecnologias digitais foram capazes de explorar esse modo de interação sem nenhuma pista da equipa do CeNTER. No entanto, levando em consideração que esse tipo de interação não é incomum e já foi utilizado com sucesso anteriormente por outras aplicações, como a “MiniCRM”<sup>3</sup> e a “Brickflow”<sup>4</sup>, existe a possibilidade de os utilizadores se familiarizarem com essa interação após o seu uso. De notar que, em termos de interação, o *swipe* entre os cartões teve uma boa aceitação entre os avaliadores, mas, de acordo com a maioria deles, não será possível usar o modo carrossel com uma grande quantidade de informações. Segundo alguns especialistas, uma alternativa poderia ser usar um formato de lista em vez de cartões.

Outra contribuição importante está relacionada com a dificuldade em definir as categorias “Atividade” e “Iniciativa”. Para a maioria dos especialistas, esses dois títulos são iguais e correspondem à mesma ideia. A equipa concordou em alterar o nome “Atividade” para “Evento” pois, no contexto do protótipo da aplicação CeNTER, a atividade está relacionada a qualquer ação ou evento que ocorra em local e data ou período específico. Para além disso, a interface foi bem avaliada pelos peritos de ambos os painéis, nomeadamente a paleta de cores usada na idealização e desenvolvimento do protótipo. O tamanho e a fonte utilizados também foram apreciados pelos especialistas dos dois painéis.

Por fim, é importante enfatizar que o conceito do protótipo da aplicação CeNTER foi bem avaliado entre os especialistas dos dois painéis. Nesse assunto, os componentes “Recursos” e “Voluntários” foram destacados como os mais inovadores da plataforma: *“Recursos e voluntários são os elementos diferenciadores da plataforma”*. (Painel 2 - Especialista 5). Como os especialistas do segundo painel estão mais familiarizados com os serviços de inovação territorial disponíveis, esse grupo de peritos provou ser uma opção valiosa para a avaliação do protótipo, pois foi capaz de contribuir significativamente para o sucesso da plataforma: *“Sugere-se converter a aplicação num website para ser acessível a mais pessoas em Portugal. Tornem-a uma alternativa mais universal, porque é uma ferramenta inovadora e pode ser complexa e exigente para o utilizador mediano.”* (Painel 2 - Especialistas 1, 5). Por outro lado, o painel 1 forneceu *inputs* técnicos essenciais, que afetam diretamente a usabilidade do protótipo: *“É necessário um feedback visual no drag and drop: animação de absorção de cartão em salvar / excluir. Passar o dedo na diagonal pode levar ao erro de arrastar a recomendação.”* (Painel 1 - Especialistas 2, 3, 4).

## 5.1.2 Avaliação Heurística

A análise das heurísticas revelou que a avaliação do protótipo foi positiva. Verificou-se que, no geral, o protótipo possui uma interface simples, o *layout* dos ecrãs apresenta-se apropriado, fácil de utilizar e não exigiu muito esforço por parte dos utilizadores. Não foram detetadas incongruências na forma de apresentação, embora alguns erros tenham sido indicados na passagem de *links* e no uso de terminologias específicas na área de Saúde e Bem-Estar.

A avaliação heurística do protótipo foi realizada através da comparação entre os resultados dos dois painéis de peritos. Para consolidar os resultados da validação heurística executados pelos dois painéis, foi criada uma

---

<sup>3</sup> <https://www.minicrm.io/tour/crm/>

<sup>4</sup> <https://brickflow.com>

grelha que reuniu todos os dados. Em seguida, as grelhas foram analisadas e comparadas. A Tabela 6 mostra o número de problemas identificados por cada grupo de avaliadores, a severidade dos problemas e a severidade média.

*Tabela 6. Número e severidade dos problemas e severidades médias identificados por cada painel de avaliadores*

Painel	Total de Problemas	Severidade dos problemas				Média de severidade
		1	2	3	4	
Peritos em tecnologias digitais	109	46	39	21	2	1,78
Peritos em turismo, saúde e bem-estar	46	14	26	6	0	1,53

Como é possível ver na Tabela 6, o painel o de especialistas em tecnologia identificou um número significativamente maior de problemas de usabilidade (109), comparativamente com o painel de especialistas em turismo, saúde e bem-estar (46). Em relação à severidade dos problemas, é importante observar que o Painel 1 identificou 46 problemas pontuados como 1 (problema estético), 39 problemas com a pontuação 2 (pequeno problema de usabilidade), 21 problemas com a pontuação 3 (problema principal de usabilidade) e 2 problemas com a pontuação 4 (catástrofe de usabilidade). Dois problemas com a pontuação 4 foram identificados pelo mesmo avaliador do painel 1. Esses problemas se referem à heurística: “Correspondência entre o sistema e o mundo real” e são relativos aos itens “As interações propostas na aplicação são semelhantes às ações reais” e “As informações aparecem em uma ordem lógica e natural”. Não foram identificados problemas com pontuação 4 no segundo painel de especialistas.

Com a maior Severidade Média, o Painel 1 identificou as heurísticas: Ajuda e Documentação (2,75), Reconhecimento em vez de memorização (2,5) e Interação entre pessoa e a aplicação (2,33). O painel 2 relatou a maior média de severidades heurísticas: flexibilidade e eficiência (2,25), ajuda e documentação (2,2), reconhecimento em vez de memorização (2) e interação entre pessoa e aplicação (2). Portanto, ambos os painéis foram bastante consistentes na avaliação da gravidade da heurística violada. A gravidade média nos dois painéis foi baixa, com uma pontuação próxima a 1,78 (Painel 1) e 1,53 (Painel 2).

A diferença nos resultados entre os painéis também ocorreu no tipo de problemas identificados. Os problemas identificados pelo Painel 1 foram mais focados no design, robustez e acessibilidade do protótipo (Tabela 7). Os problemas identificados pelo Painel 2 estavam mais relacionados com a adequação do protótipo para cada domínio de aplicação, bem como à eficiência e tipo de funcionalidade disponível para os utilizadores.

Como é possível ver na Tabela 7, as heurísticas que apresentam mais problemas de acordo com o Painel 1 incluiu: Legibilidade e *layout* (14, severidade média 1,5), seguido de controlo e liberdade (13, severidade média 1,54) e estética e design minimalista (12, gravidade média 1,66). O Painel 2 identificou o número máximo de heurísticas violadas: controlo e liberdade (11, severidade média 1,45), prevenção de erros (7, severidade média 1,86) e correspondência entre o sistema e o mundo real (6, severidade média 1,5). A identificação das heurísticas violadas facilitou a identificação e a priorização de problemas que precisam de atenção urgente antes da implantação final da aplicação.

Tabela 7. Heurísticas e severidades médias identificadas por cada painel de avaliadores

nº	Heurística	Painel 1		Painel 2	
		Número de problemas	Média de severidade	Número de problemas	Média de severidade
1	Visibilidade do estado do sistema	11	2.18	1	1
2	Correspondência entre o sistema e o mundo real	11	2.27	6	1.5
3	Controlo e liberdade	13	1.54	11	1.45
4	Consistência e padrões	7	1,57	3	1,66
5	Prevenção de erros	5	1,6	7	1.86
6	Reconhecimento em vez de memorização	4	2,5	1	2
7	Flexibilidade e eficiência de uso	7	1,85	4	2.25
8	<i>Design</i> estético e minimalista	12	1.66	1	1
9	Ajude os utilizadores a reconhecer, diagnosticar e recuperar erros	10	1,1	4	1,9
10	Ajuda e documentação	4	2,75	5	2.2
11	Interação entre pessoa e aplicação	3	2.33	1	2
12	Interação física e ergonomia	8	1,75	0	0
13	Legibilidade e <i>layout</i>	14	1,5	2	1
	<b>Total</b>	<b>109</b>	<b>1,78</b>	<b>46</b>	<b>1,53</b>

## 5.2 Teste de acessibilidade

Durante o teste de avaliação de acessibilidade remota, 74 telas do protótipo foram analisadas em um teste moderado com um especialista. O teste foi realizado com suporte de planilha. Cada questão da planilha foi dividida em pontos de atenção e oportunidades de melhoria, com uma descrição extensa para apoiar uma classificação posterior. As questões foram classificadas em prioridade e esforço de desenvolvimento. A planilha tinha um total de 42 questões separadas em 13 (31%) de alta prioridade, 11 (26%) de média prioridade e 11 (26%) de baixa prioridade. Sete questões não foram implementadas, pois só puderam ser

desenvolvidas após a fase de prototipagem do projeto. Com essa priorização e classificação de esforços, foi possível preparar uma ordem de tarefas para todos os 42 problemas (Tabela 8).

*Tabela 8: Relatório de Acessibilidade*

Prioridade	Problemas relatados	Porcentagem %
Alto	13	31
Médio	11	26
Baixo	11	26
Não implementado	7	17
<b>Total</b>	<b>42</b>	<b>100</b>

Mais de 80% dos problemas apontados pelo especialista preocupam-se em compreender os elementos da tela. O especialista apontou que alguns elementos não eram consistentes entre as telas do aplicativo ou os elementos tinham baixo contraste nas cores de primeiro e segundo plano. Outro tipo de questão dizia respeito a problemas de navegação que podem ser difíceis para pessoas com deficiência visual. Movimentos especiais com os dedos, como arrastar e soltar cartões no aplicativo, foram apontados como uma tarefa difícil para alguém com deficiência. Durante a análise da planilha, algumas referências de aplicativos externos foram usadas para configurar os elementos da tela do CeNTER. Por exemplo, o Google Maps foi usado como uma referência para a tela do mapa do aplicativo CeNTER. Com base nessa análise, todos os problemas relatados estão sendo implementados no aplicativo.

### 5.3 Teste de Microcopy

Através do teste de Microcopy foi possível obter resultados que possibilitaram o aprimoramento dos elementos textuais utilizados no protótipo.

Os principais resultados do teste de Microcopy abrangem a classificação das sugestões do especialista em diferentes categorias e a análise quantitativa dos dados. Vinte e três ajustes textuais foram sugeridos pelo especialista aos 215 elementos textuais utilizados no protótipo, sendo classificados de acordo com as categorias indicadas na Tabela 9. Algumas sugestões foram relacionadas a mais de uma categoria, originando um total de 29 itens.

*Tabela 9: Categorização das sugestões feitas pelo especialista para a Microcopy do protótipo*

Domínios	Número de sugestões	Porcentagem %
Texto difícil / ambíguo	4	13,8
Inconsistência no idioma (português de Portugal e português do Brasil)	2	6,9
Comentário relacionado a terminologia específica	3	10,3
Sugestão de vocabulário alternativo	11	37,9
Erro gramatical	6	20,7
Erro ortográfico	3	10,4
<b>Total</b>	<b>29</b>	<b>100</b>

A maioria das sugestões estava relacionada a modificações no vocabulário, incluindo alteração, remoção ou adição de conteúdo textual para substituir o original. Além disso, houve um número significativo de erros relacionados a textos gramaticais e de ambigüidade / dificuldade, indicando que os principais problemas identificados estavam relacionados à sistematização de elementos lexicais e estruturação de textos.

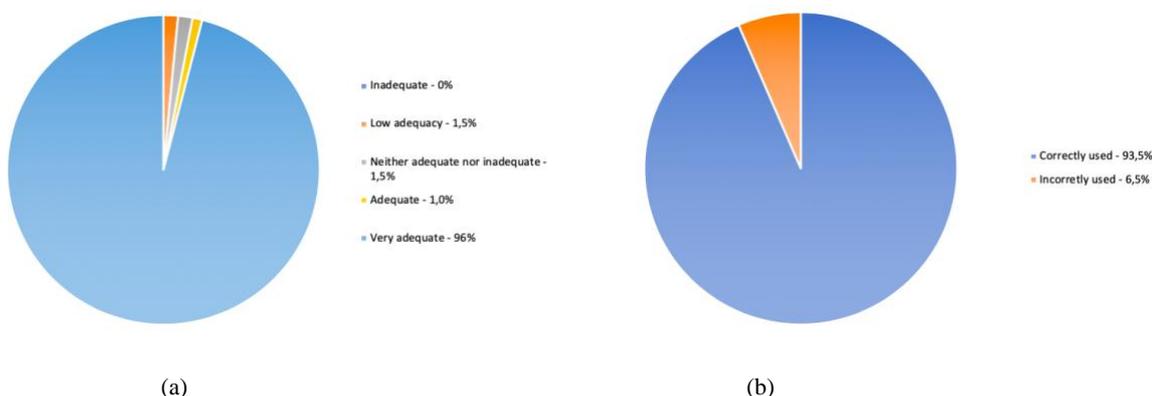
Em relação à análise quantitativa relacionada ao conteúdo correto ou incorreto, 201 de 215 (93,5%) foram classificados como usados corretamente, enquanto 14 de 215 (6,5%) foram usados incorretamente. Os resultados também são apresentados na Figura 2.

No que se refere à avaliação da adequação de cada conteúdo textual, com base em uma escala Likert de 5 pontos (a - inadequado; b - baixa adequação; c - nem inadequado nem adequado; d - adequado; e - muito adequado), o especialista pontuou cada opção na frequência a seguir (Tabela 10), que também são apresentadas na Figura 3.

*Tabela 10: A pontuação da escala Likert de 5 pontos*

Escala	0 inadequado	1 baixa adequação	2 nem adequado nem inadequado	3 adequado	4 muito adequado
Números de itens	0	3	3	2	207
Pontuação (total de 843)	0	3	6	6	828
Porcentagem	0	1,4%	1,4%	0,9%	96,3%

Levando em consideração os resultados acima, a pontuação total alcançada é de 843 (98%), o que corresponde à soma das pontuações obtidas, sendo a pontuação máxima possível de 860 (100%), comprovando que os resultados globais são positivos.



*Figure 3: (a) results of quantitative analysis related to whether the content is correct or not; (b) results concerning the assessment of the adequacy of each textual content*

Vale ressaltar que foi solicitada atenção especial aos termos relacionados aos principais conceitos do CeENTER (iniciativas, eventos, recursos, entidades, voluntários, destaques, ideias). Especificamente, em relação a esses conceitos, não foram relatados problemas pelo especialista. Portanto, os resultados apresentados acima mostram que, apesar de alguns detalhes, os conteúdos e elementos lexicais usados no protótipo CeENTER são adequados e corretamente utilizados. Os ajustes sugeridos já foram incorporados ao aplicativo.

## 6. Considerações sobre Testes com Peritos

Como o desenvolvimento do protótipo CeNTER exige um processo de design iterativo, a realização de testes em fases iniciais do desenvolvimento foi uma parte crucial no design do sistema, a fim de detectar problemas de usabilidade e melhorias importantes a serem realizadas. Através das técnicas utilizadas para a avaliação por peritos, foi possível obter dados relevantes e que deram origem a alterações substanciais no protótipo preliminar. Nesse sentido, a natureza da área de especialistas de cada painel orientou o caráter das sugestões, sendo as questões mais técnicas naturalmente apontadas pelos profissionais da área tecnológica.

Por meio dos testes realizados, o desenvolvimento de uma aplicação de inovação territorial de base comunitária mostrou-se uma proposta sustentável e oportuna, uma vez que o conceito de protótipo “CeNTER” foi positivamente avaliado. Para além disso, o protótipo apresenta uma solução inovadora pois, no ponto de vista dos especialistas, traz novas opções, como os “Recursos” e “Voluntários”, e apresenta paradigmas de interação não convencionais. Também é possível verificar uma diminuição dos valores que representam heurísticas com problemas relacionados com questões de usabilidade entre os dois painéis, de modo que é possível concluir que as alterações feitas após a primeira ronda de testes introduziram melhorias no protótipo.

Através das técnicas utilizadas para a avaliação com peritos foi possível obter dados valiosos, pois envolveram dois especialistas de diferentes áreas do conhecimento. A utilização de uma ferramenta online de avaliação de contraste permitiu identificar os problemas da tela sem intervenção humana. Os resultados deste estudo ajudaram os pesquisadores a detetar e superar os principais problemas de acessibilidade.

O teste de conteúdo de *Microcopy* permitiu verificar os termos mais adequados aos objetivos da plataforma e comprovou que os termos essenciais estão bem aplicados no contexto CeNTER. Portanto, os resultados gerais foram positivos, não exigindo melhorias complexas na plataforma CeNTER.

Este estudo proporcionou uma discussão sobre o teste de acessibilidade remota para o desenvolvimento de uma aplicação móvel, contribuindo para o conhecimento e considerando o número reduzido de estudos nesta área. A abordagem moderada para testes remotos proporcionou resultados imediatos, pois o processo de avaliação ocorreu em tempo real com a equipe. Conforme mencionado em vários estudos, esta técnica permite obter sugestões diretamente do avaliador, bem como colocar questões e obter feedback instantâneo dos avaliadores. Ao contrário, devido a um grande número de elementos de *microcopy* para avaliar, a abordagem não moderada permitiu mais tempo para analisar o conteúdo. Essa técnica garantiu, portanto, um feedback preciso. Nesse contexto, a equipe garantiu o rápido atendimento às dúvidas e esclarecimento de dúvidas do perito por telefone ou e-mail.

Algumas limitações deste trabalho incluem o fato de o *software* Principle não permitir a aplicação de algumas tecnologias assistivas, como leitores de tela, sendo justificado o auxílio de um especialista em acessibilidade. Outra restrição encontrada com o uso desse software foi a impossibilidade de recorrer a testes automatizados de conformidade para verificar a taxa de acessibilidade das fontes utilizadas e contraste de cores, embora tenha sido revisado com software que exigia entrada manual.

A validação do protótipo CeNTER não está, contudo, concluída. A próxima sessão apresenta os testes realizados com utilizadores finais. Assim, num primeiro momento serão apresentados os objetivos da avaliação, seguida do processo de operacionalização dos testes. Ainda serão apresentados os perfis de cada grupo de avaliadores, bem como os ambientes onde decorreram os testes e a versão do protótipo em vigor. Por fim, os resultados obtidos serão apresentados, e serão divididos de acordo com os respetivos casos de uso e, posteriormente, serão associados, de modo a apresentar resultados seccionais e globais. Todavia, antes de continuar os testes, as sugestões de melhorias dos especialistas foram consideradas, dando origem a uma versão mais robusta e orientada aos requisitos principais de acessibilidade, *copy*, *design* e técnicos.

## 7. Referências Bibliográficas

- Martins, A. I., Rosa, A. F., Queirós, A., Silva, A., & Rocha, N. P. (2015). European portuguese validation of the system usability scale (SUS). *Procedia Computer Science*, 67, 293-300.
- Martins, A. I., Rosa, A. F., Queirós, A., Silva, A., & Rocha, N. P. (2015). European Portuguese Validation of the System Usability Scale (SUS). *Procedia Computer Science*, 67, 293–300. <https://doi.org/10.1016/j.procs.2015.09.273>
- M. Bernhaupt, Regina; Pirker. (2013) Evaluating User Experience for Interactive Television: Towards the Development of a Domain-Specific User Experience Questionnaire, *Human-computer Interact. - Lect. Notes Comput. Sci.* 8118, 642–659.
- M. Hassenzahl, N. Tractinsky. (2006) User experience - a research agenda, *Behav. Inf. Technol.* 25, 91–97. doi:10.1080/01449290500330331.
- Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität (pp. 187–196). Vieweg+Teubner Verlag. [https://doi.org/10.1007/978-3-322-80058-9\\_19](https://doi.org/10.1007/978-3-322-80058-9_19)
- Nielsen, J. (1994a). Enhancing the explanatory power of usability heuristics. *Conference on Human Factors in Computing Systems - Proceedings*, 152–158. <https://doi.org/10.1145/259963.260333>
- Nielsen, J. (1994b) *10 Usability Heuristics for User Interface Design*. Retrieved from <http://www.nngroup.com/articles/ten-usability-heuristics>.
- Norman, D. A. (2013). *The Design of Everyday Things*. Basic Books
- Fonseca, M. J., Campos, P., & Gonçalves, D. (2012). *Introdução ao Design de Interfaces* (2a Edição). FCA - Editora de Informática. <https://www.fca.pt/pt/catalogo/informatica/design-multimedia/introducao-ao-design-de-interfaces/>
- Ballard, B. (2007). *Designing the Mobile User Experience*. Wiley & Sons, Ltd.
- Knight, W. (2019). *UX for Developers*. Northampton: Apress.
- Still, B; Crane, K. (2017). *Fundamentals of User-centered Design*. Nova York: CRC Press.
- Brennand, C. V. L. T., Hornung, H., Baranauskas, C. (2018). *Análise de Métodos de Avaliação da Experiência de Usuário* (Report No. IC-18-18). Retirado de <https://repositorio.ucs.br/xmlui/handle/11338/1340>.
- Barbosa, A. (2019, outubro 23). Medindo a usabilidade do seu produto com System Usability Scale (SUS). Retirado de <https://medium.com/design-contaazul/medindo-a-usabilidade-do-seu-produto-com-system-usability-scale-sus-3956612d9229>.



## Apêndice 1 - Grelha de validação heurística

### Grelha de validação heurística do protótipo da aplicação CeNTER

0-Não existe consenso que seja um problema de usabilidade; 1-Problema estético apenas; 2-Problema de usabilidade menor; 3-Problema de usabilidade importante; 4-Catástrofe de usabilidade

nº	Heurística	Indicadores	0	1	2	3	4	Observações
1	Tornar o estado do sistema visível	1.1 O utilizador é informado sobre o que está a acontecer na aplicação por meio de feedback. (Ex.: “Mensagem enviada”).						
		1.2 Os itens selecionados são claramente distintos dos outros.						
		1.3 As mensagens sobre o status da aplicação possuem uma linguagem clara e concisa.						
2	Correspondência entre o sistema e o mundo real	2.1 As interações propostas na aplicação são semelhantes a ações reais. (Ex: arrastar um item para o lixo).						
		2.2 O significado de símbolos e ícones são compreensíveis e intuitivos.						
		2.3 A informação aparece numa ordem lógica e natural. (Ex.: as listas de opção são ordenadas alfabeticamente).						

		2.4 As cores utilizadas correspondem às expectativas comuns sobre o código de cores. (Ex.: a cor vermelha é usada para alertar sobre problemas).							
3	Utilizador controla e exerce livre arbítrio	3.1 O utilizador pode guardar uma tarefa a meio da sua execução para a continuar futuramente. (Ex.: salvar um formulário parcialmente preenchido).							
		3.2 A aplicação oferece mensagens de conclusão para tarefas que envolvam um conjunto de etapas.							
		3.3 É possível voltar ao ecrã anterior a qualquer momento.							
		3.4 É possível sair do ecrã atual a qualquer momento.							
		3.5 Informações relevantes (Ex.: botões, comandos e textos) ficam visíveis quando o teclado é aberto.							
		3.6 A aplicação deixa claro qual o próximo passo a executar para completar a tarefa em curso.							
4	Consistência e adesão a normas	4.1 A navegação da aplicação e o <i>layout</i> dos ecrãs são consistentes. (Ex.: menus, botões de confirmação e áreas de texto principais sempre aparecem na mesma posição em todos os ecrãs).							
		4.2 Os botões distinguem-se do restante <i>layout</i> , deixando evidente que são clicáveis.							
		4.3 O nome do botão/link é consistente com o nome do ecrã/página que abre.							
5	Evitar erros	5.1 Nas primeiras interações do utilizador com a aplicação são mostradas instruções básicas.							

		5.2 Os botões e controlos que desencadeiam ações irreversíveis ficam localizados em áreas de difícil acesso e exigem uma confirmação do utilizador.						
6	Reconhecimento em vez de lembrança	6.1 Há padronização de cores para identificação e sinalização das áreas da aplicação.						
		6.2 A aplicação utiliza, uma linguagem clara e conhecida pelo utilizador.						
		6.3. Os campos que requerem a inserção de dados são facilmente identificados.						
7	Flexibilidade e eficiência	7.1 A página inicial mostra todas as opções relevantes.						
		7.2 A aplicação oferece mecanismos para uma navegação dinâmica.						
		7.3 A aplicação disponibiliza formas alternativas de efetuar a mesma tarefa.						
8	Estética e design minimalista	8.1 O design visual é atraente.						
		8.2 Os ecrãs apresentam somente informação e componentes importantes.						
		8.3 A aplicação exibe quantidades pequenas de informações em cada ecrã.						
9	Ajudar os utilizadores a reconhecer, diagnosticar e a recuperar erros	9.1 A aplicação permite fácil recuperação de erros.						
		9.2 As mensagens de erro são claras e precisas.						
		9.3 As mensagens indicam a causa do erro.						

		9.4 As mensagens indicam como corrigir o erro.						
10	Ajuda e documentação	10.1 A aplicação possui uma opção de “Ajuda” bem visível.						
		10.2 A ajuda é de fácil compreensão.						
		10.3. A aplicação facilita o regresso à tarefa depois de consultar a opção “Ajuda”.						
11	Interação entre a pessoa e a aplicação	11.1 A navegação da aplicação é intuitiva.						
		11.2 A aplicação facilita a inserção manual de informação (Ex.: opções em forma de lista, recurso de auto-completar, etc).						
		11.3 Em situações de pesquisa, a aplicação possui um mecanismo de inserção de dados tolerante a erros de digitação e uso de abreviaturas.						
12	Interação física e ergonomia	12.1 Possui botões/links com tamanho adequado ao toque.						
		12.2 Possui espaço suficiente entre botões e links, evitando que o utilizador selecione outro elemento.						
		12.3 Os botões e controles podem ser facilmente acedidos com qualquer uma das mãos.						
13	Legibilidade e layout	13.1 O espaçamento entre linhas utilizado favorece a leitura.						
		13.2 As fontes utilizadas favorecem a leitura.						
		13.3. A interface da aplicação adapta-se a rotação do ecrã.						
		13.4 Os ícones/textos possuem contraste suficiente em relação ao plano de fundo.						

		13.5. As imagens possuem cor e detalhes favoráveis à leitura.							
		13.6. A aplicação realça os conteúdos mais importantes.							

Data \_\_\_\_\_

Local \_\_\_\_\_

Avaliador nº \_\_\_\_\_