

Testes com utilizadores

Turno IPM201113L08, Grupo 17A

Daniel Correia, 80967 | Carolina Inês Xavier, 81172 | Inês Leite, 81328

Guião dos testes

Objetivo - O que se pretende atingir com o teste?

Pretendemos com estes testes verificar como é que os utilizadores realizam as tarefas, quais as dificuldades encontradas, os tempos necessários para realizar as mesmas, e os aspetos positivos da interface, de modo a poder melhorá-la para interações futuras.

Onde – Onde são realizados os testes?

Todos os testes serão realizados nas mesmas condições, podendo o ambiente variar.

Quando – Quando serão os testes?

Os testes serão realizados entre os dias 13.05.2016 e 16.05.2016.

Duração – Quanto tempo durará cada sessão de testes?

A duração máxima prevista para uma sessão de avaliação é de 20 minutos.

Equipamento – Qual o equipamento informático necessário para realizar os testes?

Vamos precisar de um computador e um telemóvel (cronómetro) para realizar os testes.

Software – Qual o software necessário para o teste?

Para a realização dos testes, o software necessário é a aplicação por nós desenvolvida.

Estado - Qual deve ser o estado do nosso sistema no início do teste?

Os utilizadores vão começar o teste no ecrã inicial.

Tempo de resposta - Qual deve ser a carga e tempo de resposta do sistema onde corre a nossa aplicação?

A aplicação será usada quando o sistema está em carga normal.

Coordenador e Observador – Quem irá conduzir e observar os testes?

Coordenador: Carolina Xavier

Observadores: Daniel Correia, Inês Leite

Utilizadores – Quem serão os utilizadores que farão os testes e quantos serão necessários?

Os testes serão realizados por 17 utilizadores principiantes. 3 dos utilizadores serão mulheres e os restantes, 14, homens, com idades compreendidas entre os 16 e os 56 anos. Todos os utilizadores frequentam bares.

Tarefas – Que tarefas serão pedidas aos utilizadores?

Serão realizadas as seguintes tarefas, de forma aleatória para não viciar certos resultados:

- Tarefa 1 - Pedir uma vodka com gelo e duas águas e pagar logo com multibanco;
- Tarefa 2 - Adicionar a música “Kool Thing” à lista de reprodução, fazer like nessa música e partilhar a música atual no Facebook com o nome de utilizador: Pedro e palavra-passe: 123;
- Tarefa 3 - Chamar um táxi para avenida D João II e cancelá-lo de seguida.

Fim correto – Que critério será usado para determinar o fim da execução de uma tarefa corretamente?

As 3 tarefas estão terminadas quando o utilizador chega ao menu onde está a resposta.

Ajuda – Que ajudas (manuais, ajuda, online, etc) estarão disponíveis para o utilizador durante os testes?

Não serão dadas quaisquer tipos de ajuda sob a forma de manual em papel ou ajuda online.

Ajudar - Em que medida se pode ajudar os utilizadores durante os testes?

Só podemos ajudar os utilizadores quando estes bloqueiam em algum passo durante a realização de uma tarefa, e caso não se recordem da tarefa a realizar.

Dados – Que dados irão ser recolhidos e como serão analisados?

Será recolhido o tempo necessário para a realização de cada tarefa por parte dos utilizadores, bem como os erros por eles cometidos ao realizar cada tarefa. Será analisado o tempo das tarefas, e se está

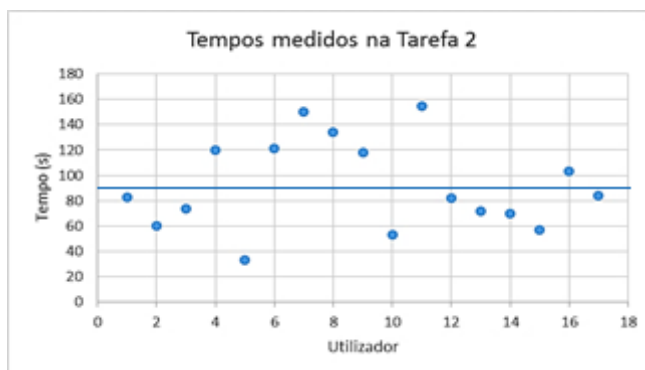
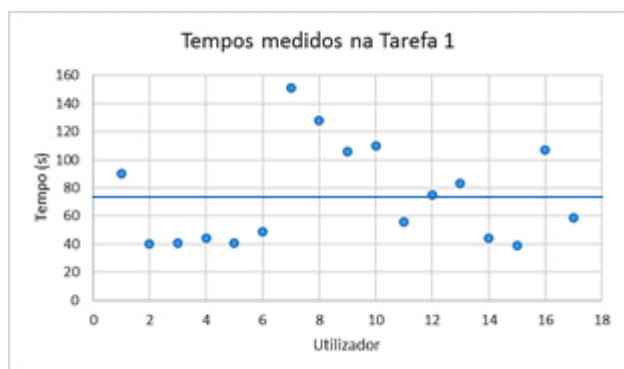
dentro dos intervalos de tempo esperado, e possíveis causas para os erros cometidos pelos utilizadores durante a realização das tarefas.

Sucesso – Qual o critério de sucesso que ditará que a interface é um sucesso?

A nossa interface será considerada um sucesso se cumprir todos os objetivos iniciais.

Medições dos tempos para executar as tarefas

Tarefas	Média	Desvio-padrão	Intervalo de Confiança
Tarefa 1	74	35	[57,91]
Tarefa 2	92	14	[75,109]
Tarefa 3	30	14	[23,37]



Análise do nível de satisfação dos utilizadores:

Dificuldade	Média	Desvio-padrão	Mediana	Moda
Tarefa 1	3	0,80	2	2
Tarefa 2	3	0,93	3	2
Tarefa 3	2	0,51	2	2

Sendo a Escala de Likert utilizada para avaliar a Satisfação:

1 – Muito fácil, 2 – Fácil, 3 – Normal, 4 – Difícil, 5 – Muito difícil.

Medições dos erros cometidos:

Tarefa 1:

- Dificuldade em encontrar o separador de bebidas não alcoólicas.
Taxa de frequência do erro: 41%
- Dificuldade em encontrar o botão “personalizar” na lista de pedidos.
Taxa de frequência do erro: 24%
- Pesquisar pela bebida “água” no separador de bebidas alcoólicas.
Taxa de frequência do erro: 47%

Tarefa 2:

- Carregar na música em vez de carregar no ícone (+) para adicionar música à lista de reprodução.
Taxa de frequência do erro: 12%
- Dificuldade em encontrar o botão ‘Adicionar música à lista de reprodução’.
Taxa de frequência do erro: 41%
- Dificuldade em encontrar a página correta onde fazer *like* na música.
Taxa de frequência do erro: 24%
- Não perceber qual a música atual. O enunciado da tarefa provocou o erro.
Taxa de frequência do erro: 6%

Tarefa 3:

- Carregar na opção “Não” ao chamar um táxi, para o cancelar o táxi. O enunciado da tarefa provocou o erro.
Taxa de frequência do erro: 6%

Comparação com critérios de usabilidade:

Tarefas	Tempo previsto	Tempo médio
Tarefa 1	< 120 segundos	74
Tarefa 2	< 120 segundos	92
Tarefa 3	< 60 segundos	30

Eficiência

- Analisando a tabela acima, podemos concluir que os resultados obtidos se encontram dentro dos critérios estipulados para estes testes, tendo a tarefa 2 o tempo médio mais elevado.

Satisfação

- Em relação aos critérios de Satisfação, os resultados obtidos para a tarefa 2 e 3 estão dentro do intervalo esperado. No caso da tarefa 1, a classificação média foi 3 o que está fora do intervalo definido.

Eficácia

- As tarefas 1 e 2 foram completas com sucesso por 82% dos utilizadores. A tarefa 3 foi completa com sucesso por todos os utilizadores.

Tarefa	Sucesso (%)
Tarefa 1	82
Tarefa 2	82
Tarefa 3	100

Conclusões sobre cada tarefa

- Na tarefa 1 o tempo médio efetuado pelos utilizadores é mais baixo que o estipulado. Apesar de não estarmos a espera muitos utilizadores tiveram dificuldades em encontrar o separador das bebidas não alcoólicas demorando mais tempo a executar a subtarefa “Pedir duas águas”. O botão de “personalizar” não foi intuitivo para os utilizadores.
- Na tarefa 2 apesar de o tempo médio ser inferior ao tempo previsto, o número de erros cometidos pelos utilizadores foi bastante alto. Este acontecimento deve-se ao facto do botão de adicionar música à lista de reprodução e o ícone (+) serem pouco perceptíveis, criando confusão em alguns utilizadores, que não sabiam onde clicar durante a execução da tarefa.
- Na tarefa 3 o tempo médio efetuado pelos utilizadores é metade do tempo previsto. Esta é também a tarefa onde os utilizadores cometeram menos erros.

Conclusão geral

No geral a nossa aplicação passou aos testes com sucesso, os resultados obtidos não diferem muito do esperado. Conclui-se assim que as funcionalidades de pedidos de bebidas/comida e de música podem ser melhoradas, consoante sugestão dos utilizadores, enquanto que a funcionalidade dos transportes não parece necessitar de qualquer melhoria.