

Introdução às Interfaces Pessoa-Máquina

Laboratório 11 – Grupo 4

Instituto Superior Técnico – 2017/2018

Grupo 4: Bernardo Valente, Francisco Machado, Pedro Teixeira.

Turno quinta-feira 11h30

Guião de teste com utilizadores

Preparação

Para a realização da avaliação o avaliador deve ter consigo um cronómetro, caneta e um papel para anotar medidas de desempenho e eventuais comentários/observações do utilizador. O equipamento necessário, para o utilizador, consiste apenas num computador.

As tarefas devem ser realizadas num ambiente calmo e silencioso evitando distrações por parte do utilizador.

O utilizador deve preencher um pré-questionário para recolha de dados demográficos e um formulário de consentimento, confirmando que os dados recolhidos podem ser utilizados em trabalhos relacionados com a cadeira de IIPM.

Introdução

O protótipo iRave que irá ser apresentado foi desenvolvido para a cadeira de Introdução às Interfaces Pessoa-Máquina. É um dispositivo *wearable* que presta funções úteis a frequentadores de festivais de música.

O objetivo de testar os utilizadores é avaliar a interação dos mesmos, com o protótipo. Será pedido ao utilizador que realize algumas tarefas usando o dispositivo.

A recolha de dados da interação dos utilizadores com o iRave permitirá averiguar erros de funcionalidade do protótipo durante a execução de cada tarefa e apurar aspetos a melhorar.

Para que toda a recolha de dados seja feita de forma legal o utilizador deve dar autorização para que a informação recolhida seja utilizada em trabalhos futuros para a cadeira de IIPM. Esta será utilizada para otimização e melhoramento do protótipo.

Pedimos agora que o utilizador preencha um pequeno pré-questionário para recolha de dados demográficos. Estes dados serão usados apenas para avaliação estatística.

O Pré-Questionário é composto pelas seguintes perguntas:

- 1 - Género
- 2 - Faixa Etária
- 3 - Familiarização com smartwatches
- 4 - Familiarização com computadores
- 5 - Frequência de participação em festivais
- 6 - Ocupação

Metodologia

A avaliação das tarefas deve ser precedida do preenchimento do pré-questionário entregue aos utilizadores, para recolha de dados demográficos. Os utilizadores executam as tarefas em ordem aleatória num computador portátil usando o *trackpad*.

Cada tarefa começa no ecrã de bloqueio e é cronometrada a partir do momento em que o utilizador faz o primeiro clique. Devem ser cronometrados e registados os tempos que cada utilizador demora a cumprir cada tarefa, tal como o número de erros que faz. Opiniões e informações relevantes partilhadas pelos utilizadores devem ser, igualmente, registadas.

Uma tarefa termina assim que a última instrução for cumprida. Os erros cometidos pelos utilizadores devem ser registados pelo avaliador. Um erro é cometido se o utilizador escolher uma opção que não o direcione para o ecrã suposto.

No fim de cada tarefa o utilizador tem de a classificar a tarefa de acordo com Single Ease Question, ou seja, numa escala de 1 a 7, sendo 1 Muito Difícil e 7 Muito Fácil, a dificuldade da tarefa executada.

Avaliação

Os utilizadores devem realizar as seguintes tarefas:

- **Tarefa 1:** partilhar a sua localização com o amigo *Feli*.

- **Eficácia:**
 - *Medida:* número de erros.
 - *Critério:* em média, menos de 1 erro.
- **Eficiência:**
 - *Medida:* tempo de execução da tarefa.
 - *Critério:* em média, menos de 30 segundos.
- **Satisfação:**
 - *Medida:* Single Ease Question.
 - *Critério:* em média, resultados superiores a 5 pontos.

- **Tarefa 2:** Adicionar à lista de favoritos o artista *Frank Ocean*, que toca no dia 1, e consultar a lista de favoritos.

- **Eficácia:**
 - *Medida:* número de erros.
 - *Critério:* em média, menos de 2 erros.
- **Eficiência:**
 - *Medida:* tempo de execução da tarefa.
 - *Critério:* em média, menos de 40 segundos.
- **Satisfação:**
 - *Medida:* Single Ease Question.
 - *Critério:* em média, resultados superiores a 4 pontos.

- **Tarefa 3:** Adicionar à lista de favoritos todos os artistas que vão atuar no palco 2 e consultar a lista de favoritos.

- **Eficácia:**
 - *Medida:* número de erros.
 - *Critério:* em média, menos de 2 erros.
- **Eficiência:**

- *Medida*: tempo de execução da tarefa.
- *Critério*: em média, menos de 50 segundos.
- **Satisfação**:
 - *Medida*: Single Ease Question.
 - *Critério*: em média, resultados superiores a 4 pontos.

Balanço

Para terminar a avaliação por favor refira os pontos fortes e os pontos fracos da aplicação.
O grupo agradece tempo despendido na realização destas tarefas, todos os comentários são muito importantes para o desenvolvimento da aplicação.

Experiência

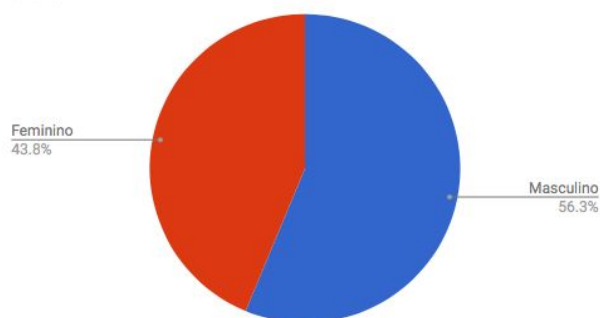
Os utilizadores foram avaliados individualmente pelos três elementos do grupo. A escolha dos utilizadores foi feita de forma aleatória mas tentou-se incidir na faixa etária de pessoas que frequentam mais regularmente festivais de música.

Cada avaliação foi executada individualmente e sem interrupções. A ordem pela qual se executavam as tarefas foi completamente aleatória. Os utilizadores efetuaram a avaliação em computadores portáteis usando o *touchpad* dos mesmos.

Características dos Participantes:

Foram inquiridas 16 pessoas com a seguinte distribuição de idades e sexo:

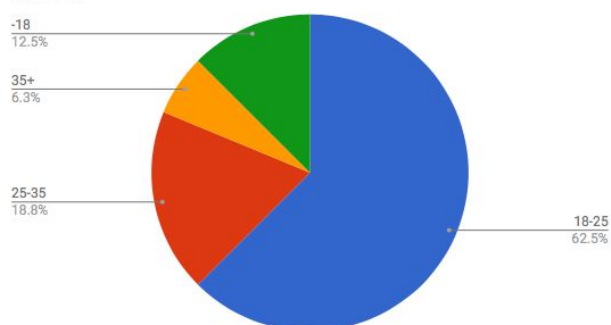
Sexo



Das 16 pessoas inquiridas, 11 são estudantes e 43,8% são do sexo feminino, como podemos observar no gráfico.

Os restantes utilizadores têm como ocupações trabalhos como professor, advogados e motoristas. Note-se que todos estes encargos têm uma relação com tecnologias *smartwatch* ou outras quase nula.

Idades



Analisando os dados demográficos recolhidos podemos concluir que grande parte da amostra escolhida está relativamente bem familiarizado com tecnologias *smartwatch*. As exceções são pessoas que não pertencem à população alvo do iRave. Relativamente à familiarização dos inquiridos com o uso de computadores portáteis, todos os sabiam operar.

Análise dos Resultados

Tarefas e métricas de usabilidade esperadas:

- **Tarefa 1:** partilhar a sua localização com o amigo Feli.
Objetivos de Usabilidade: Em média, os utilizadores devem demorar cerca de 30 segundos a executar esta tarefa. Cometem no máximo um erro e as pessoas consideram esta tarefa de fácil e rápida resolução.
- **Tarefa 2:** adicionar à lista de favoritos o artista Frank Ocean e de seguida consultar a lista de favoritos.
Objetivos de Usabilidade: Em média, os utilizadores devem demorar cerca de 40 segundos a executar esta tarefa. Cometem no máximo dois erros e as pessoas consideram esta tarefa acessível.
- **Tarefa 3:** adicionar à lista de favoritos todos os artistas que vão atuar no palco 2 e de seguida consultar a lista de favoritos.
Objetivos de Usabilidade: Em média, os utilizadores devem demorar cerca de 50 segundos a executar esta tarefa. Cometem no máximo dois erro e as pessoas consideram esta tarefa acessível.

Foram observados as seguintes medidas de desempenho:

- Tempo para completar uma tarefa;
- Número de erros cometidos;
- Grau de dificuldade de cada tarefa.

Primeiro comparámos os valores obtidos para os erros e duração de cada tarefa com os métricas de usabilidades esperadas.

	Objetivo		Média (obtida)		Desvio Padrão (obtido)	
	Erros	Duração	Erros	Duração	Erros	Duração
Tarefa 1	<1	~30 segundos	0,43	18,04 segundos	0,72	5,82 segundos
Tarefa 2	<2	~40 segundos	0,93	39,71 segundos	1,23	14,4 segundos
Tarefa 3	<2	~50 segundos	0,81	43,44 segundos	0,81	15,14 segundos

Como os valores obtidos do desvio padrão são elevados é conveniente recorrer a outro tipo de forma estatística . Utilizámos o Intervalo de Confiança (IC) como recomendado nas aulas, com $p = 0,05$. Para o cálculo do IC usa-se para além do parâmetro p , o desvio e o número de dados analisados.

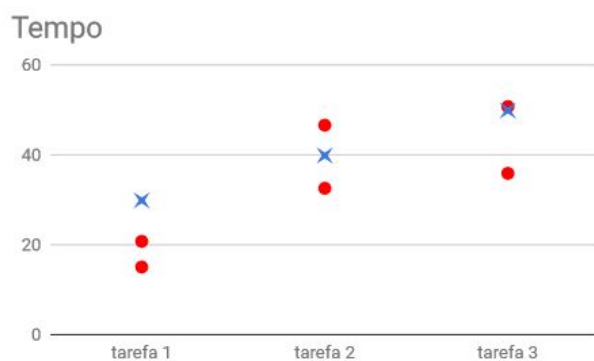
Intervalo de confiança

$$x_{\min} = \text{Média} - C$$

$$x_{\max} = \text{Média} + C$$

Parâmetro C calculou-se usando a operação CONFIDENCE no programa Excel.

Para os tempos obtiveram-se os seguintes intervalos de confiança:



Tarefa 1: Esperado: <30

Xmin = 15,19

Xmáx = 20,9

Tarefa 2: Esperado: <40

Xmin = 32,7

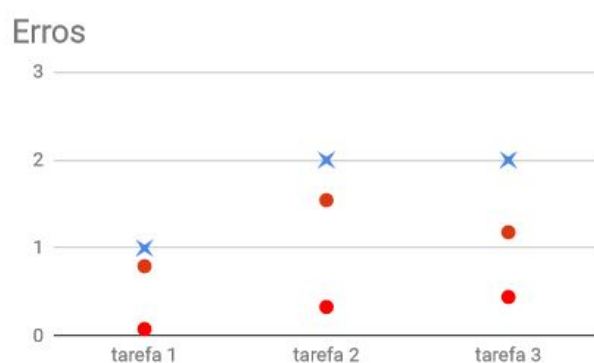
Xmáx = 46,7

Tarefa 3: Esperado: <50

Xmin = 36,0

Xmáx = 50,86

Para os erros obtiveram-se os seguintes intervalos de confiança:



Tarefa 1: Esperado: <1

Xmin = 0,08

Xmáx = 0,79

Tarefa 2: Esperado: <2

Xmin = 0,33

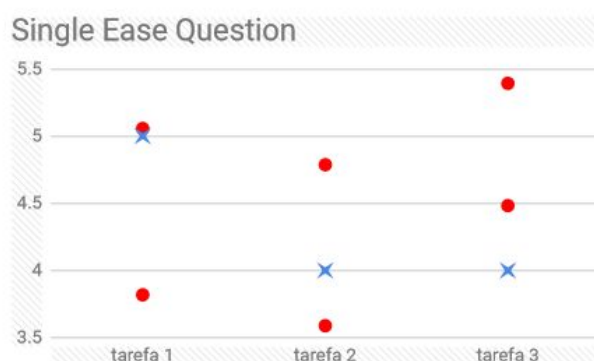
Xmáx = 1,54

Tarefa 3: Esperado: <2

Xmin = 0,44

Xmáx = 1,17

Para os resultado de Single Ease Question obtiveram-se os seguintes intervalos de confiança:



Tarefa 1: Esperado: >5

Xmin = 3,82

Xmáx = 5,06

Tarefa 2: Esperado: >4

Xmin = 3,59

Xmáx = 4,79

Tarefa 3: Esperado: >4

Xmin = 4,48

Xmáx = 5,39

Resultados:

Resultado esperado?	Tarefa 1	Tarefa 2	Tarefa 3
Eficiência	Sim	Não	Não
Eficácia	Sim	Sim	Sim

Satisfação	Não	Não	Sim
------------	-----	-----	-----

Discussão Global

Existe alguma discrepância nos resultados esperados e os obtidos.

Como visto há tarefas cuja usabilidade não é assegurada. Esta situação pode dever-se ao facto de não se terem realizados desenvolvimentos intermédios antes de efectuado o protótipo final. Por exemplo, uma maneira de poder ter este tempo objetivo assegurado seria ter-se realizado mais testes com utilizadores numa fase mais precoce do protótipo tendo assim acesso a informação que contribuiria para a realização das tarefas no tempo esperado.

Em relação à tarefa 1 que passa no critério de usabilidade tanto na eficiência como na eficácia poderá dever-se ao facto de se tratar de uma função mais simples, no entanto o resultado dos utilizadores quanto à satisfação não foi o esperado.