

Avaliação Comparativa da Usabilidade de Locais na World-Wide Web

3.5

Um botão descriptivo "notiço"

Daniel J Gonçalves, João Brisson Lopes, Joaquim A Jorge, Mário Rui Gomes

Departamento de Engenharia Informática

Instituto Superior Técnico / Universidade Técnica de Lisboa

djvlg@gia.ist.utl.pt, brisson@ist.utl.pt, jorgej@acm.org, mrg@ist.utl.pt

recomendações? design gráfico (artístico)

Sumário

A extraordinária difusão da World-Wide Web e protocolos associados nestes últimos anos, transformaram os navegadores Internet e páginas Web numa das interfaces utilizador mais visíveis no mundo depois do telefone e do automóvel. Este artigo apresenta um estudo prático comparativo da usabilidade de locais noticiosos e de venda de livros, duas das actividades mais populares entre a comunidade *on-line*. Através deste estudo conduzido no âmbito de uma disciplina de licenciatura do IST, procurámos caracterizar a facilidade de utilização de várias abordagens distintas, nacionais e estrangeiras numa experiência de âmbito limitado (quatro locais e 130 utilizadores). A aplicação de técnicas de avaliação heurística mesmo com base num questionário simples e limitado permitiu identificar uma surpreendente quantidade de problemas de usabilidade nalguns dos locais estudados, mesmo naqueles com excelente reputação internacional.

Palavras-chave

Usabilidade, Avaliação Heurística, WWW, Experiência de Utilização.

1. INTRODUÇÃO

A avaliação heurística [Nielsen94] consiste num excelente instrumento para identificar problemas de usabilidade numa interface utilizador por forma a corrigi-los numa abordagem iterativa. Este tipo de avaliação, quando utilizado criteriosamente permite identificar problemas significativos que afectam a maioria das interfaces. No caso da World-Wide Web, os sistemas de informação encontram-se obrigatoriamente estruturados em termos do paradigma de hipertexto que realça naturalmente os aspectos ligados à navegação. A metáfora espacial consiste num princípio de organização forte determinando a nossa percepção do meio, reflectindo-se nas características que mais frequentemente procuramos na rede: liberdade de movimentos, caminhos claramente assinalados, acesso rápido e, sobretudo, respostas imediatas.

Claramente os atributos desejáveis de uma interface utilizador estão presentes na WWW, embora as características específicas deste meio nos levem a avaliar a experiência de utilização em termos da facilidade de acesso e, em última análise, da navegabilidade dos locais.

O objectivo do trabalho que apresentamos, inscrito no currículo prático de uma disciplina de licenciatura do IST, consistiu na sensibilização dos utilizadores para os factores de usabilidade na World-Wide Web, numa análise comparativa com outras abordagens mais convencionais. A WWW apresenta novos desafios e refor-

mula antigas equações no sentido de os tempos de resposta extremamente elevados relativamente aos sistemas interactivos convencionais provocarem alterações tanto nas expectativas como no comportamento dos utilizadores. Cremos que o presente estudo realça alguns destes factores. Por outro lado, através de um questionário simples procurou-se estabelecer a ligação entre características de desenho e problemas de usabilidade.

De notar que neste estudo introdutório não se faz uma cobertura exhaustiva das principais heurísticas de usabilidade mas apenas das mais notórias. Nielsen [Nielsen98] aponta dez características fundamentais de usabilidade das quais salientamos algumas das mais relevantes para o presente estudo:

Apresentar o estado do sistema de forma clara e visível. Este é fundamentalmente um critério de navegação. As duas questões mais importantes que os utilizadores se colocam em ambientes hipertextuais são: "Onde estou?" e "Para onde posso ir em seguida?". Os locais *devem* ter cada página claramente assinalada e com indicação da secção a que pertence.

O controle e a liberdade de escolha pertencem ao utilizador. Embora a maior parte dos navegadores ofereçam um conjunto substancial de ajudas de navegação (incluindo o botão Back, sem dúvida o artefacto de navegação mais utilizado [Nielsen94]), ter uma estrutura de navegação clara e consistente em cada página constitui um meio eficaz de evitar a desorientação dos utili-

zadores. Mecanismos simples, como a presença de um botão Home em cada página, são eficazes ao auxiliar os utilizadores a recuperarem de erros de navegação.

Desenho consistente e aplicação de normas. Os utilizadores não deveriam ser confundidos perante diferentes significados ou acções diferentes para a mesma palavra. Uma forma de evitar isto é adoptar convenções consistentes para navegação, etiquetas de botões, títulos de páginas, legendas em elos hipertexto e *seguir-las de modo consistente* em todas as páginas de um local.

Mais vale prevenir erros do que remediá-los. Enquanto não existirem bibliotecas de *Widgets* de utilização difundida, o *JavaScript* pode ser utilizado para verificar erros triviais no preenchimento de formulários. Por outro lado, toda a atenção dada à navegação do local é pouca. A informação de navegação é importante para prevenir erros ao seleccionar a página de destino.

Ênfase no reconhecimento em vez da memorização. Um bom desenho permite tornar objectos e acções visíveis. Os utilizadores não deveriam ser obrigados a memorizar informação entre várias componentes de um diálogo. As instruções para o uso do sistema deveriam ser claras e bem visíveis. Se os utilizadores reconhecerem onde estão ao examinarem a página corrente, têm menos oportunidades de se perderem. Uma boa navegação deve fornecer mapas para ajudar na tarefa de orientação.

Flexibilidade e eficiência de uso. Providenciar "aceleradores" para facultar acesso rápido a várias funções do sistema é uma prática louvável. Exemplos desta funcionalidade são as marcas ou *bookmarks* suportadas pelos navegadores. As páginas do local devem ser desenhadas para serem marcáveis ou acessíveis através de elos. Estes critérios tornam mais delicado o uso de quadros (*frames*), e requerem atenção no caso dos resultados de buscas.

Desenho estético e minimalista. Os diálogos não deveriam conter informação irrelevante. Animações supérfluas são um dos exemplos mais gritantes da violação deste princípio. Outro problema mais subtil é a inclusão de informação excessiva numa página. Demasiado texto/imagens distraem os utilizadores, sendo preferível um desenho que através de elos judiciosos fornece a informação pretendida em níveis de detalhe crescentes sem sobrecarregar cada diálogo (e o utilizador) com informação irrelevante.

Com estes critérios em mente desenhámos uma experiência para comparar bons desenhos com realizações menos conseguidas centrando o método experimental em métricas de tempo de execução e erros incorridos ao executar tarefas simples.

2. A EXPERIÊNCIA

O protocolo experimental estabelecido para a experiência especificava as tarefas a realizar, as regras a observar durante a realização das tarefas e como proceder ao registo das apreciações.

As tarefas a realizar incidiram sobre quatro locais na WWW, sendo dois locais noticiosos e os outros dois duas livrarias com venda via WWW. Os locais noticiosos escolhidos foram a Agência Lusa e a CNN. As livrarias foram a Amazon e a Livraria Arco-Íris. Apenas foram dados aos utilizadores os nomes dos locais a visitar, cabendo-lhes encontrar os respectivos URL, numa tentativa de, por meio de exemplos, demonstrar pedagogicamente a necessidade de os URL terem correspondência lógica com os nomes das respectivas organizações.

Nos locais noticiosos, as tarefas consistiram em, através de consulta dos locais da experiência, tentar obter resposta às seguintes questões:

1. Em que data será lançada no Japão a nova consola *Playstation 2* ?
2. Que programas aconselha a CNN para *streaming* de áudio e vídeo na sua página ?
3. Que deveria escrever, para fazer uma pesquisa, no site da Agência Lusa, de artigos sobre consolas que não refiram a *Dreamcast* ?

Para as livrarias, as quatro tarefas a realizar pelos utilizadores consistiram em consultar os locais das livrarias para responder a:

1. Descreva, sucintamente, quais as modalidades de envio e pagamento.
2. Quanto custa a obra "Maré Alta Estelar" (*Startide Rising*), de David Brin?
3. Quantos livros de Isaac Asimov foram publicados na Colecção Nébula (Arco-Íris) e pela editora Bantam (Amazon)?
4. "Encomende" a obra "As Brumas de Avalon" (*The Mists of Avalon*), de Marion Zimmer Bradley (seguir todo o processo até à introdução de dados pessoais).

3. PROCEDIMENTO E REGRAS

Como os grupos de laboratório são constituídos por dois alunos, decidiu-se realizar as respeitando esta disposição, o que tem a vantagem adicional de reflectir a realidade de muitos grupos de trabalho na maioria das empresas de pequena dimensão. Em tais grupos é difícil estabelecer procedimentos que garantam a separação total entre observador e observado e estes não comuniquem entre si, mas é possível tirar partido desta aparente limitação [Buur99]. Se o observador seguir um protocolo de "voz alta", vocalizando as suas acções e sentimentos, existirá um aumento da quantidade de informação que pode ser observada e o observador poderá abandonar a postura clássica em que não intervém (e, portanto, não influencia os resultados, [Dumas93]) e emitir sugestões ou estímulos cujas consequências, positivas ou negativas, constituem uma fonte de informação adicional. Este procedimento é muito importante quando, por exemplo, se pretenda elucidar causas de insucesso ou bloqueio do utilizador no desempenho das tarefas.

Os dois elementos de cada grupo foram também instruídos para alternarem os papéis de observador e observado entre tarefas quando procedessem à execução destas.

Ao observador do grupo competia registar o tempo que o utilizador observado necessitava para completar cada tarefa, as reacções demonstradas e/ou vocalizadas e o número de erros cometidos. Dado a avaliação directa desta última medida ser algo problemática, estabeleceu-se que se deveria entender como de erro qualquer das situações seguintes:

1. Erros de **Navegação**, incluindo o retorno (através do botão *Back*) de elos hipertextuais conducentes a páginas não desejadas ou com informação sem utilidade para completar a tarefa.
2. Páginas com **carregamento interrompido pelo utilizador**, quer ao abandonar uma página antes de terminar o seu carregamento quer ao interrompê-lo através de tecla de função ou de menu contextual.
3. **Erros no preenchimento de campos de um formulário**. Estes erros são geralmente devidos a uma má concepção dos formulários. O erro no preenchimento de campos deve-se muitas vezes a má informação ou insuficiência dos textos de apoio ou ainda à formulação incorrecta das perguntas.
4. **Pesquisas que não produzem os resultados pretendidos**. A pesquisa permite formas de navegação mais sofisticadas que o atravessamento de elos hipertextuais. No entanto o preenchimento incorrecto de formulários ou a incompreensão pelos utilizadores da forma mais correcta de obter resultados evidenciam de um modo geral erros de concepção do local cuja correcção é imperativa.

Finalmente, o protocolo explicitava que, após visitarem todos os locais constantes da experiência, os elementos do grupo deveriam produzir uma apreciação subjectiva segundo seis critérios indicados nos formulários de registo, concertando opiniões entre observador e utilizador, de acordo com o recomendado por [Buur99] para equipas de pequena dimensão.

Esta apreciação subjectiva feita por mútuo acordo entre os elementos de cada equipa, contemplou seis aspectos diferentes. (1) *A facilidade de navegação*, (2) *Aspecto visual das páginas*, (3) *Legibilidade*, incluindo utilização judiciosa de tipos de letra e disposição geral dos textos, (4) *Utilização de cores* e sua contribuição para apresentação da informação, (5) *Utilização de animações* e (6) *Apreciação global* sobre a qualidade e usabilidade do local.

Cada item deveria ser classificado numa escala de 0 a 3, sendo o número de valores intencionalmente par com vista a obviar o conhecido fenómeno de escolha do valor médio por omissão. Não se explicitaram quaisquer critérios para a métrica e deixou-se esta tarefa aos utilizadores.. Este desvio intencional dos procedimentos correctos tinha por objectivo pedagógico demonstrar as consequências do estabelecimento de métricas incompletas nos resultados experimentais, nomeadamente a grande variabilidade subjectiva existente no estabeleci-

mento de tais critérios. Este assunto será retomado mais adiante quando analisarmos os resultados obtidos.

Nas secções seguintes resumimos os principais comentários produzidos pelos alunos para cada local avaliado.

3.1 Agência Lusa <<http://www.lusa.pt>>

O local está aparentemente mais destinado a profissionais da informação, utilizando organização e termos não familiares para utilizadores comuns. Isto conduz a dificuldades na navegação, nomeadamente na utilização do motor de procura, resultantes por exemplo, da necessidade de especificar o intervalo de tempo a que respeitam as notícias. Pelo que verificámos, os estudantes normalmente não possuem uma noção clara de quando ocorrerem os factos das notícias, causando dificuldades à pesquisa.



Figura 1. Página de Entrada da Agência Lusa

Existe uma grande zona na página de entrada destinada a uma animação com o logotipo da agência, visível apenas com um navegador *específico* (*Internet Explorer, versão 4*). A maioria dos utilizadores ao aceder com um navegador diferente, deparou-se com uma área em branco, criticando o facto ao afirmar que este prejudica o aspecto visual da página de entrada.

Muitas das críticas visaram o aspecto gráfico geral. Embora considerando a cor de fundo adequada e a legibilidade bastante boa, as restantes cores foram usadas em excesso e sem fim aparente. O aspecto das páginas foi considerado de fraca concepção e pouco atractivo e as animações desnecessárias, monótonas e cansativas. Este aspecto ilustra uma falha comum a muitos locais que procuram atrair a atenção dos utilizadores com animações que, na realidade, conseguem apenas distraí-los das tarefas que os atraíram inicialmente ao local (predominantemente a busca de informação, no caso dos locais noticiosos). A pouca informação existente na página principal foi diversas vezes elogiada. Apesar disso, a organização da página, bem como um excesso de publicidade, foram severamente criticados, nomeadamente quanto à disposição dos diversos elementos, sem ordem aparente e dando mais preponderância a aspectos pouco relevantes (contador de acessos, etc.) do que às notícias em si (grave, dada a natureza do local).



Figura 2. Página Principal da CNN

A taxonomia usada nos temas foi considerada inadequada e pouco descritiva dos conteúdos, ilustrando a importância de falar a linguagem do utilizador, aderência a normas e as virtudes de um *design estético e minimalista*, que não relevam no caso vertente.

3.2 CNN <<http://CNN.com>>

O *design* deste local está aparentemente mais vocacionado para o público em geral, não surgindo os problemas de falta de familiaridade encontrados com a Agência Lusa. A boa qualidade da página é consensual. O local aparece repetidamente descrito como agradável, atractivo e com boa apresentação nas avaliações. As cores foram usadas harmoniosamente e realçam o essencial. Pela positiva destacam-se ainda o motor de procura e a utilização correcta de animações. De notar que estas estão geralmente associadas à publicidade, dado consistirem num meio de atrair fortemente a atenção dos utilizadores. A sua utilização excessiva distrai e irrita os navegantes, levando a um aumento do número de erros na realização das tarefas.

O principal aspecto negativo referido foi a excessiva quantidade de informação, sobretudo na página principal, provocando tempos de carregamento longos e dificultando a navegação. Por exemplo, o acesso às versões em língua estrangeira do local, bem como à função de procura, é difícil dado as opções correspondentes estarem diluídas no meio da informação noticiosa. Vários utilizadores criticaram a legibilidade, prejudicada pelo tamanho dos caracteres, e a utilização das cores na página principal notando o uso excessivo muitas cores em áreas pequenas conduzindo a uma concepção com falta de consistência.

Os avaliadores sugeriram diminuir a quantidade de informação por página (no que diz respeito à barra de navegação e toda a parafernália que se segue à notícia principal) e uma estruturação das notícias, (ou a inclusão de tópicos como "o que?", "como?", "quando?", etc. associados a cada uma delas) para permitir um fácil acesso à informação relevante. De notar que a quantidade de informação presente nas páginas de locais situados nos EUA exacerba os mais pequenos problemas de usabilidade no contexto em que a experiência foi efectuada

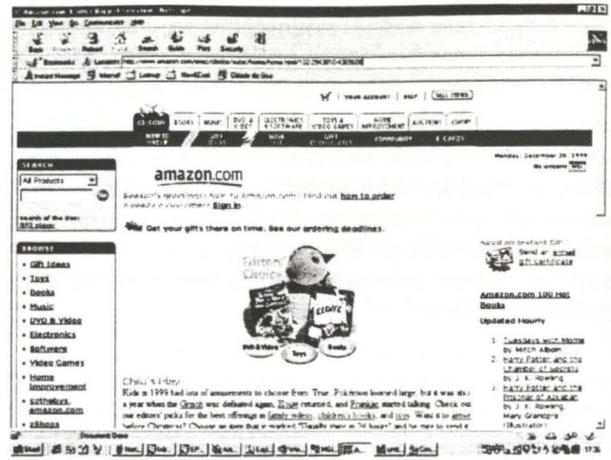


Figura 3. Página Principal da Amazon

ada (laboratórios do IST), devido principalmente à saturação das ligações para o exterior, nomeadamente aos EUA.

3.3 Amazon <<http://www.amazon.com>>

A Amazon foi escolhida por se tratar de um dos locais mais conhecidos e bem sucedidos a nível da World-Wide Web. A sua concepção estética e despojada, em comum com outros locais de elevado tráfego, ilustra uma das características de usabilidade na WWW. A navegação mereceu os comentários mais positivos dos alunos que destacaram a versatilidade e flexibilidade dos mecanismos de orientação, disponibilizando várias formas de completar a mesma tarefa. Por outro lado, a utilização de cores foi também positivamente apreciada pelo esquema de cores uniforme e combinação agradável. Quanto à legibilidade, porém, as opiniões dividiram-se, tendo as avaliações subjectivas variado entre as notas mínima e máxima.

Apesar da lentidão no carregamento das páginas referida por muitos utilizadores, o local foi considerado muito bom ou excelente. A metáfora do carrinho de compras foi elogiada pela simplicidade da representação. Mesmo assim, algumas avaliações apontaram que, para quem não conhece o sistema, este pode ser confuso, embora de aprendizagem rápida.

O aspecto mais criticado do local foi que, apesar de conter informação de boa qualidade, as páginas são densas, com demasiada informação apresentada ao mesmo tempo. Esta observação reflecte sem dúvida a recente diversificação do modelo de negócio da Amazon, que passou a incluir música, jogos e aparelhagens electrónicas além de leilões de bens de consumo, numa tentativa de não ser ultrapassada pelos concorrentes. Um exemplo disto é a referência feita por vários alunos à existência de muitos elos para escolher numa dada página, nem sempre sendo claro onde conduzem.

Algumas das sugestões de melhoramentos apontaram para estruturar a informação resultante das pesquisas, como a subdivisão por línguas, por exemplo.

3.4 Livraria Arco-Íris <<http://www.liv-arcoiris.pt>>

Este local foi objecto de muitas críticas dos utilizadores desde a inutilidade da página de entrada, considerada uma perda de tempo, ao aspecto visual considerado pela maioria dos alunos como mal conseguido e mal arrumado, denotando um *design* amadorístico. A grande maioria achou que as cores foram usadas em número excessivo e não adequadamente. O local foi considerado de difícil leitura devido à utilização de letras pequenas e informação concentrada. Por outro lado, os utilizadores reflectiram também críticas à má organização do texto, caracterizada como dispersa e confusa. Outros aspectos negativos indicados foram as animações agressivas, como por exemplo uma intitulada "veja destaques" e a quantidade excessiva de gráficos. Pela positiva destacou-se a existência de um menu permanente facilitando a navegação, apesar de a colocação da barra de navegação do lado direito e não do lado esquerdo (mais comum) ter sido mencionada.



Figura 4. Página principal da livraria Arco-Íris

O motor de pesquisa utiliza uma base de dados já disponibilizada anteriormente pela livraria. Em termos funcionais, este foi não só considerado limitado (só se podendo aceder à pesquisa avançada depois de uma busca simples sem sucesso), como a base de dados foi considerada bastante incompleta.

Em suma, o local foi considerado pouco intuitivo, em termos de navegação e aprendizagem, obrigando os navegantes a avançar por tentativa e erro até conseguirem usá-lo com alguma eficácia devido à desorientação causada por múltiplas fontes de distração. De ressaltar que é difícil comparar em termos equitativos o orçamento gasto pela Amazon, cujo movimento de negócios se cifra em centenas de milhões de contos, com os gastos diminutos envolvidos na concepção do local da livraria Arco-Íris.

4. RESULTADOS

A avaliação foi efectuada por 57 grupos de dois alunos em oito turnos de duas horas entre 28 de Setembro e 2 de Outubro de 1999. Em cada turno a avaliação foi con-

duzida em simultâneo por seis a dez grupos utilizando outros tantos computadores pessoais Pentium II Celeron 266 MHz com 4 Gb Disco, Placa Som, Monitor 17", operando sob Windows NT e utilizando os dois navegadores mais populares.

Aos utilizadores foi pedido para produzirem as suas avaliações após o exame dos locais e atribuída uma quota de 20 minutos dedicáveis a cada local. Não houve preocupação em visitar os locais por ordem arbitrária, o que pode, de alguma forma, ter polarizado os resultados. No final da avaliação de cada local os estudantes preencheram inquéritos que entregaram na semana subsequente após recomendarem o redesenho de um dos locais.

4.1 Taxa de Sucesso

Considerou-se como resultados válidos as tarefas com uma duração não superior a 10 minutos, o qual segundo a nossa experiência representa o limite superior do tempo que o utilizador está disposto a esperar. Segundo este critério, obteve-se uma taxa de sucesso para cada tarefa cujos valores são apresentados na Tabela 1. Verifica-se que as taxas de menor sucesso correspondem a tarefas em que foi necessário realizar consultas mais complexas a motores de procura.

Verifica-se igualmente que, no caso das livrarias, a taxa de sucesso no local da Livraria Arco-Íris é inferior à do local da Amazon quando a tarefa implica procura de preços e encomenda de livros, demonstrando a melhor adequação do local da Amazon à tarefa fundamental deste locais, a venda de livros.

A Tabela 2 compara a média do número de erros cometidos quando a tarefa foi realizada em tempo inferior ao tempo limite (sucesso) e a média correspondente quando o tempo da tarefa excedeu esse limite (insucesso) para cada uma das tarefas realizadas.

O nível de significância, isto é, a probabilidade de duas médias correspondentes serem semelhantes, é também apresentado na tabela. Um valor baixo deste nível (as médias não são estatisticamente semelhantes) significa que os utilizadores não completaram a respectiva tarefa por se terem perdido nos locais, não tendo encontrado a página ou páginas com a informação pretendida. Os outros casos poderão prenunciar outros problemas encontrados pelos utilizadores tais como grande degradação do tempo de resposta da rede ou dos próprios locais.

Tarefa	Local			
	CNN	LUSA	Amazon	Arco-Íris
1	74%	93%	86%	98%
2	89%		96%	91%
3		96%	74%	95%
4			94%	84%

Tabela 1 – Percentagem de sucesso nas tarefas.

Local	CNN		LUSA		Amazon		Arco-Íris	
	Média	Signif.	Média	Signif.	Média	Signif.	Média	Signif.
1 Sucessos	1,05	0%	2,57	4%	1,31	92%	0,53	0%
	Insucessos	5,60	13,25		1,38		8,00	
2 Sucessos	0,88	2%			0,15	53%	0,54	31%
	Insucessos	3,00			1,50		3,20	
3 Sucessos			1,52	17%	1,88	0%	0,87	30%
	Insucessos		6,50		4,43		1,33	
4 Sucessos					0,63	19%	2,02	3%
	Insucessos				2,33		7,56	

Tabela 2 – Número de erros de navegação cometidos por tarefa.

4.2 Medidas de Usabilidade

As medidas de usabilidade realizadas consistiram na determinação do número de erros cometidos pelos utilizadores ao realizarem as tarefas e a duração das mesmas. A Tabela 3 apresenta as correspondentes médias e os percentis 10 e 90 das tarefas realizadas com sucesso.

Da análise dos resultados obtidos ressalta imediatamente que os tempos necessários à conclusão de uma tarefa são, em geral, mais elevados nos locais situados nos EUA do que nos locais portugueses. As exceções são as tarefas 2 e 4 realizadas nos locais Amazon e Arco-Íris em que os tempos apresentam médias (e dispersões) com uma semelhança provável de 90 % no caso da tarefa 2 e de 70% no caso da tarefa 4. Este resultado reforça a ideia (ver análise da taxa de sucesso) de que o local da Amazon teve cuidado especial nas tarefas de procura de preços e de realização de encomendas.

No tocante ao número de erros de navegação, verifica-se significativa diferença do local da Lusa em relação à CNN. Os utilizadores notaram este facto e atribuíram-no à dificuldade em usar o motor de procura do local da Lusa, principalmente à forma desorganizada da respectiva interface. Ainda sob este aspecto, na comparação entre os locais da Amazon e da Arco-Íris, verifica-se uma situação semelhante à do tempo para completar as tarefas. Cometem-se mais erros no local da Amazon ao realizar as tarefas 1 e 3 e mais erros no local da Arco-Íris ao realizar as tarefas 2 e 4. O maior cuidado posto pelo local da Amazon nas tarefas centrais do negócio (compras) dá origem a menos erros que compensam o tempo de resposta mais longo, na duração verificada destas tarefas.

O tempo necessário para concluir uma operação como medida de usabilidade aparece distorcido pelos tempos de espera forçados pelo comportamento não determinístico da rede. Este comportamento exacerba ainda alguns dos problemas de usabilidade de locais como a CNN que colocam uma quantidade excessiva de informação por página de acordo com os julgamentos da maior parte dos utilizadores. Tais problemas não surgem com ligações de alta velocidade. No entanto, mes-

mo nos EUA, estudos realizados [GVU98] revelam que cerca de 45% dos utilizadores acedem à Internet através de ligações de baixa velocidade (33Kbps ou inferior), o que torna o tempo de carregamento das páginas um factor dominante nas transações.

4.3 Avaliação Subjectiva

A Tabela 4 apresenta os resultados da apreciação subjectiva dos locais visitados pelos utilizadores ao empregarem critérios de avaliação heurística.

Desde logo é possível constatar que, em geral, a apreciação dos locais estrangeiros é melhor do que a dos locais portugueses. A exceção mais significativa verifica-se na legibilidade dos textos dos locais noticiosos, em que o local da CNN, devido à sua densidade informativa e uso de tipos reduzidos, foi penalizado pelos utilizadores, e considerado como semelhante ao local da Agência Lusa cujo *design* é bastante pobre.

Por outro lado o critério de utilização excessiva de animações não foi entendido desta forma pelos utilizadores que em grande parte atribuíram classificações pelo número de animações presente num local em vez da sua utilização excessiva, um critério estético de avaliação. Este facto realça a importância de se tornar bem claros os critérios de valoração nos questionários. Embora as pontuações apareçam artificialmente baixas, a avaliação final favorece os locais com uso mais "discreto" (ou inexistente) de animações.

De um modo geral a taxa de erros e de tarefas concluídas acompanha as apreciações subjectivas dos utilizadores sobre a qualidade dos locais. Acreditamos que estes factores em conjunto com o número de acções de navegação, número de elos seguidos e buscas sejam métricas fiáveis de usabilidade. Na generalidade, as apreciações favoráveis sobre facilidade de navegação e qualidade geral dos locais concordam com as observações sobre a taxa de erros e operações concluídas com êxito.

Das outras medidas heurísticas de usabilidade, destacamos a legibilidade dos textos, em que a CNN com grande densidade informativa e tipos de tamanho reduzido foi penalizada pelos utilizadores. A Amazon foi a melhor classificada de acordo com este critério.

Local	CNN			LUSA			Amazon			Arco-Íris			
	Média	Percentil 10	Percentil 90	Média	Percentil 10	Percentil 90	Média	Percentil 10	Percentil 90	Média	Percentil 10	Percentil 90	
1	Erros	1,05	0,00	2,90	2,57	0,00	6,00	1,31	0,00	3,00	0,53	0,00	2,00
	Duração	5,74	3,00	9,00	4,07	1,00	7,80	6,09	2,35	10,00	2,79	1,00	5,16
2	Erros	0,88	0,00	2,00				0,15	0,00	0,00	0,54	0,00	2,00
	Duração	3,94	1,00	8,00				2,65	1,00	5,00	2,70	1,00	5,40
3	Erros				1,52	0,00	5,10	1,88	0,00	4,00	0,87	0,00	2,70
	Duração				3,43	1,00	7,10	6,53	3,00	10,00	3,74	1,00	7,70
4	Erros							0,63	0,00	2,20	2,02	0,00	3,80
	Duração							5,85	3,00	10,00	5,64	2,00	8,80

Tabela 3 – Médias e percentis 10 e 90 dos erros de navegação e tempos das tarefas realizadas com sucesso

5. APRECIÇÃO GLOBAL

Os resultados obtidos apresentam, em geral, uma grande dispersão dos tempos de execução das tarefas tal como se pode ver nas Figuras 5 e 6 através dos histogramas de distribuição das medidas em duas tarefas típicas: procura de uma notícia e encomenda de um livro. A origem desta dispersão pode ser encontrada em duas variáveis independentes: o tempo de resposta da rede e a variabilidade das características individuais dos utilizadores, dado que estes, apesar de pertencerem a um grupo aparentemente homogéneo à partida (idade, orientação profissional), são na realidade pessoas muito diferentes.

A variação do tempo de resposta da rede provocaria variabilidade na execução das tarefas de turno para turno consoante as condições verificadas no dia e hora a que a experiência teve lugar, enquanto a variabilidade das características pessoais provocaria a existência de dispersão em cada turno. A análise dos resultados obtidos ilustra claramente este fenómeno, o que prefigura que esta segunda variável independente seja preponderante nos resultados. No entanto, os dados obtidos não permitem eliminar a influência do tempo de resposta da rede, antes permitem afirmar que a sua influência é menor do que a da variabilidade das características pessoais dos utilizadores.

Um ponto comum às anotações dos utilizadores respeita aos motores de procura. A experiência obrigou os alunos ao emprego sucessivo de quatro motores de busca diferentes, tanto na funcionalidade como na interface utilizador. Isto constitui um factor adicional de confusão verificando-se que os utilizadores, inconscientemente, tentaram aplicar aos motores de procura dos locais examinados por último modelos mentais adquiridos durante o uso dos primeiros. O resultado traduziu-se no insucesso da procura e na frustração que, em alguns casos, levou ao abandono das tarefas. Estes factos foram exacerbados quando não existiam auxílios adequados que explicassem concisamente, como as buscas deveriam ser realizadas.

Outro facto negativo consiste nos motores de pesquisa não permitirem manter o contexto, ou seja, permitirem uma procura sobre o conjunto resultante de outra. Esta observação mostra que os utilizadores desta experiência tinham conhecimentos de bases de dados.

Uma constatação importante desta avaliação tem a ver com o estabelecimento das métricas a empregar e com o significado que os utilizadores lhes atribuem. A métrica empregue nesta experiência estabeleceu uma escala de 0 a 3 para todos os critérios apreciados, mas não explicitou qualquer critério para a sua atribuição. Daqui resultaram tantas interpretações do significado dos valores da escala quantos os grupos intervenientes na experiência, normalmente atribuídas casuisticamente, tal como os seguintes exemplos mostram:

Alguns alunos atribuíram o valor 2 ao item "Legibilidade" com a justificação "É muito confusa".

Outros grupos atribuíram valor 1 ao item "Utilização de Animações" com a justificação "Utilizam poucas animações" para um local onde não existem animações e ao qual obviamente este critério não seria aplicável.

Outros alunos usaram a classificação 1 no item "Utilização de Cores" justificando-o com o comentário "Muito mau" a que deveria corresponder o valor 0 (note-se a relutância a atribuir o valor 0).

Alguns utilizadores tentaram estabelecer critérios formais, embora de forma incompleta e aplicados inconsistentemente.

Estes factos evidenciam a pouca fiabilidade a atribuir a inquéritos com questões que se limitam a pedir uma classificação numa escala numérica sem explicitação de quaisquer critérios e mostram a conveniência em evitar tais escalas, substituindo-as pelos critérios, ainda que, posteriormente, as respostas sejam processadas por meio de valores numéricos. O anterior mostra ainda que os critérios devem ser estabelecidos e impostos externamente.

Local	CNN			LUSA			Amazon			Arco-Íris		
	Média	Percentil 10	Percentil 90	Média	Percentil 10	Percentil 90	Média	Percentil 10	Percentil 90	Média	Percentil 10	Percentil 90
Facilidade de navegação	2,09	1,00	3,00	1,62	1,00	3,00	2,17	1,00	3,00	1,95	1,00	3,00
Aspecto visual das pags.	2,14	1,00	3,00	1,37	1,00	2,40	2,40	2,00	3,00	1,65	1,00	2,30
Legibilidade	1,86	1,00	3,00	1,83	1,00	3,00	2,18	1,00	3,00	1,71	1,00	3,00
Utilização de cores	2,31	2,00	3,00	1,54	1,00	2,40	2,52	2,00	3,00	1,42	0,60	2,00
Utilização de animações	1,77	0,50	3,00	1,12	0,00	2,00	1,55	0,00	3,00	1,38	1,00	2,00
Apreciação global	2,10	1,00	3,00	1,34	1,00	2,00	2,41	2,00	3,00	1,70	1,00	2,00

Tabela 4 – Médias e percentis 10 e 90 das avaliações heurísticas dos quatro locais da experiência

Por outro lado, as opiniões dos utilizadores dividem-se quanto à legibilidade de um mesmo texto. A par de comentários negativos como "Legibilidade 1, muito fraca devido à densidade e muita informação", outros utilizadores comentam sobre a legibilidade do mesmo texto como sendo "Boa", atribuindo-lhe o valor máximo da escala (3). O tamanho dos caracteres é pois característico de preferências pessoais. Verifica-se ainda que a maioria dos utilizadores desconhece a possibilidade de modificar o tamanho por omissão do tipo de texto oferecida pela grande maioria dos navegadores.

Destaca-se igualmente a pouca importância que os locais portugueses dão à utilização de vários navegadores e o facto do arranjo das páginas não realçar o conteúdo. A disposição das páginas não leva em linha de conta que os ecrãs possam apresentar resoluções diferentes do ambiente em que as páginas foram criadas e testadas. De assinalar também a inexistência de um critério homogéneo quanto às áreas de colocação de elos e de formulários de busca, dificultando a navegação.

O comportamento observado enquanto os utilizadores tentavam cumprir as tarefas que lhes foram atribuídas concorda com as suas críticas posteriores aos locais. Com excepção dos tempos de carregamento (por vezes demasiado lentos), as tarefas a efectuar nos locais da CNN e da Amazon decorreram de forma simples. Como excepção, a pesquisa dos vários livros de Isaac Asimov publicados pela Bantam não foi, de um modo geral imediata, obrigando várias vezes a refinar os critérios de pesquisa. Mais importante, nas tarefas relacionadas com a aquisição de livros, o local da Amazon demonstrou estar concebido para a execução expedita destas tarefas.

Notou-se alguma diferença entre os utilizadores mais familiarizados com a WWW (e alguns até com os locais navegados, que já tinham usado anteriormente) e os restantes. No entanto, a boa funcionalidade dos locais da CNN e da Amazon minimizou de forma significativa essa diferença. Quanto ao local da Agência Lusa, embora as tarefas tenham sido realizadas de forma relativamente simples, embora não imediata, o sentimento geral, manifestado sob a forma de vários comentários, foi

de insatisfação, principalmente no que diz respeito à forma de efectuar as procuras. Neste local nota-se a orientação para o público com formação jornalística, perito em certas modalidades de pesquisa noticiosa, em claro detrimento do público em geral.

Finalmente, o local da Arco-Íris mostrou-se o de utilização mais difícil. Muitos grupos não conseguiram cumprir todas as tarefas, o que se revelou frustrante e contraproducente para o fim último da loja: a venda de livros.

6. CONCLUSÕES

Quando iniciámos o trabalho reportado neste artigo o principal objectivo que pretendíamos atingir era o de sensibilizar os alunos para a importância do desenho na experiência de utilização. A esta fase e após o ensino teórico de regras e princípios a usar no desenho e o ensino prático das potencialidades de linguagens como o HTML e o JavaScript, pretendíamos efectuar uma experiência de avaliação heurística mais detalhada listando mais factores de usabilidade, a qual está, à data da escrita deste artigo, em fase de execução.

Após a execução de todos os testes e análise dos melhoramentos propostos pelos vários grupos a aspectos específicos dos locais, achamos que a experiência devia ser reportada. Assim o trabalho procurou ilustrar, de uma forma simplificada, as virtudes e limitações da avaliação de locais na World-Wide Web na execução de tarefas realistas. Este estudo permitiu ainda revelar a importância mas também as dificuldades práticas de avaliação de interfaces na Web que constitui um meio fundamentalmente diferente dos sistemas interactivos tradicionais.

A primeira conclusão a retirar do trabalho efectuado é a importância da avaliação de locais, mesmo dos mais prestigiados a nível internacional. Estes locais apresentam, nas suas páginas, uma quantidade de informação demasiado grande cujo carregamento só é viável através de linhas de comunicação de alto débito, o que ainda não é uma realidade nem sequer nos Estados Unidos. Assim as medidas de desempenho, em termos de tempo

de execução de tarefas, nem sempre constituem indicadores fiáveis de usabilidade ao discordarem da avaliação subjectiva através de critérios heurísticos.

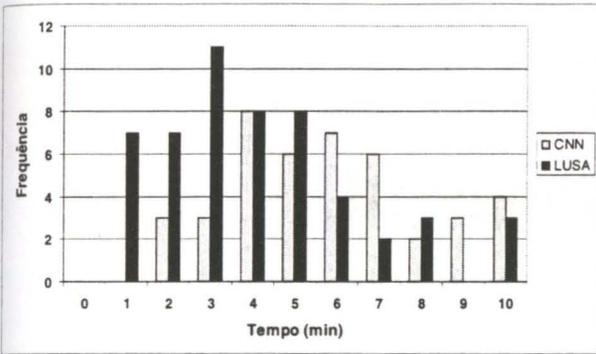


Figura 5 – Distribuição do tempo de procura de uma notícia.

Uma segunda conclusão importante é a de que um desenho correcto dos questionários assume grande importância na avaliação final, ao permitir a comparação directa das respostas individuais.

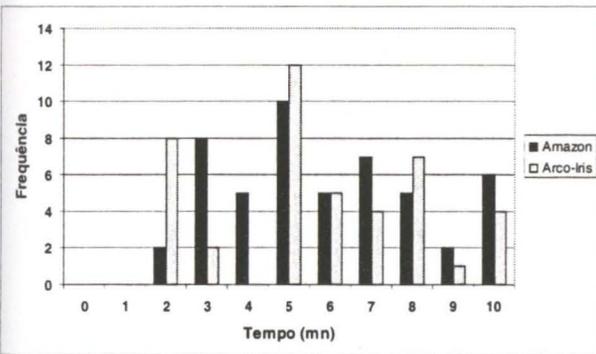


Figura 6 – Distribuição do tempo de encomenda de um livro.

O presente trabalho mostrou ser necessário investigar a importância das variáveis-independentes como o tempo de resposta da rede, características e atitude dos utilizadores nos resultados das medidas de usabilidade clássi-

cas, tais como o tempo para completar a tarefa, número de erros e elos hipertextuais seguidos.

Estamos certos que a experiência reportada neste artigo se revelou muito importante para uma melhor avaliação de locais e, principalmente, para que os nossos alunos adquiram os conhecimentos necessários ao desenho de interfaces utilizáveis para além de critérios meramente estéticos.

7. BIBLIOGRAFIA

- [Buur99] Buur, J., Bagger, K., Replacing Usability Testing with User Dialogue, Communications ACM, vol. 42, n.º 5, pp. 63-66, May 1999.
- [Dumas93] Dumas, J., Redish, J., A Practical Guide to Usability Testing, Ablex Publishing Corp., 1993.
- [GVU98] Gvu's 10th WWW User Survey, Georgia Tech Graphics, Visualization and Usability Center, 1998.
- [IHM99] Documentação da disciplina de Interfaces Utilizador, Semestre de Outono de 1999 <http://www.rnl.ist.utl.pt/~ic-ihm/>.
- [Nielsen94] Jakob Nielsen, Alert Box, 1994 <http://www.useit.com/alertbox/>.
- [Nielsen98] Jakob Nielsen, Usability Metrics: How good are You?, DevHead, ZDNet Publications, Agosto 1999. <http://www.zdnet.com/devhead/stories/articles/0,4413,2321306,00.html>
- [Preece93] Jenny Preece (ed.), A Guide to Usability: Human Factors in Computing, Addison-Wesley, 1993.
- [Rosenfeld98] Rosenfeld, Information Architecture for the World Wide Web, O'Reilly & Associates, Mar 1998.
- [Spool98] Spool, Jared M et al, Web Site Usability: A Designer's Guide, Morgan Kaufmann Publishers, Nov 1998.