

7 Experimentos Práticos

Existem diversas metodologias possíveis para a avaliação de uma interface, tais como observação e monitoração de usuários, coleta da opinião dos usuários, experimentos ou testes de *benchmark*, avaliação interpretativa e avaliação preditiva (Preece et al., 1994). Neste trabalho, optou-se por realizar experimentos com usuários, pois esta metodologia utiliza uma perspectiva semi-científica e permite uma avaliação mais controlada dos resultados. Em comparação com os demais métodos citados, os experimentos permitem que os dados sejam coletados e analisados quantitativamente em situações mais específicas. Entretanto, devido à complexidade das interações entre as pessoas e a interface, a validade dos resultados obtidos muitas vezes pode ser questionada.

Este capítulo apresenta testes de usabilidade do protótipo de interface assistida para deficientes visuais desenvolvido neste trabalho. Apesar de ter sido empregadas metodologias de IHC para avaliação de interfaces, não houve intenção de aplicar formalmente os métodos de Engenharia de Usabilidade (Nielsen, 1993) nem realizar uma análise teórica profunda dos resultados obtidos. Os experimentos servirão de base para uma avaliação preliminar das hipóteses sobre a solução dos problemas apresentados anteriormente e também mostrar que protótipo implementado é útil para os deficientes visuais.

7.1 Seleção de Usuários

A Tabela 7.1 ilustra o perfil dos usuários selecionados para os testes. Todos estão na faixa etária de 15 a 38 anos. O sexo e o grau de instrução não foram relevantes nem para a seleção e nem para a análise quantitativa dos resultados. Com relação à percepção visual, foi intencionalmente selecionado um maior número de usuários completamente ou praticamente cegos e buscou-se mesclar cegos desde o nascimento e também cegos há poucos ou vários anos. A classificação adotada para medir a percepção visual (veja coluna correspondente

na tabela) é a mesma utilizada nas práticas esportivas, conforme apresentada no Capítulo 2.

Usuário	Idade	Sexo	Percepção Visual	Cego há (anos)	Escolaridade	Utiliza PC?	Categoria
01	38	M	B3	Nasc.	1º	+/-	Reab.
02	29	M	B1	5	2º	Não	Reab.
03	17	M	B2	10	1º	+/-	Aluno
04	23	F	B2	Nasc.	1º	Não	Aluno
05	24	F	B1	Nasc.	1º	Não	Aluno
06	18	M	B2	14	1º	Não	Atleta
07	18	F	B1	9	1º	+/-	Aluno
08	20	M	B1	17	2º	Sim	Atleta
09	38	M	B1	20	3º	Sim	Atleta
10	26	M	B1	Nasc.	2º	Sim	Atleta
11	29	M	B1	5	1º	Não	Reab.
12	15	M	B1	12	1º	Sim	Aluno
13	18	M	B1	12	1º	+/-	Aluno
14	20	M	B1	Nasc.	2º	Sim	Atleta
15	38	F	B1	3	2º	Não	Reab.

Tabela 7.1: Usuários Selecionados

Cerca de 50% usuários selecionados têm como causa principal da perda de visão os problemas de catarata, glaucoma, retinose pigmentar ou descolamento de retina. Cerca de 20% apresentam problemas associados ao nervo ótico, tais como inflamações ou atrofias. Cerca de 30% apresentam outros problemas ou não souberam identificar o motivo pelo qual perderam a visão.

Para escolher os usuários, foram feitas algumas visitas ao Instituto Benjamin Constant. Em cada uma, um grupo específico de usuários em potencial foi reunido

e os objetivos e idéias principais do trabalho foram apresentados em linguagem simples. As pessoas puderam também experimentar livremente um pouco o equipamento (*joystick* reativo), porém sem nenhum compromisso. Foi feita uma rápida entrevista e elaborado um cadastro com todos os interessados em participar dos testes. Todos os cadastrados foram, portanto, voluntários. Entretanto, devido à perspectiva de utilizar uma tecnologia nova e capaz de facilitar o acesso a computadores, notou-se um grande grau de interesse e motivação em quase todas as pessoas.

Os grupos específicos mencionados anteriormente incluem três categorias distintas (veja coluna correspondente na tabela): alunos, atletas e reabilitandos. Os alunos freqüentam os cursos regulares do Benjamin Constant, ou seja, o equivalente ao primário (primeiro grau) das escolas convencionais. Os atletas praticam algum esporte na instituição, geralmente por lazer, tais como futebol ou corrida. Finalmente, os reabilitandos são pessoas que estão buscando o desenvolvimento de sua visão residual (quando há) e dos demais sentidos. Eles freqüentam cursos de mobilidade e orientação espacial (onde aprendem, por exemplo, a utilizar bengalas para caminhar), aulas de natação, dança, violão, massagem, trabalhos manuais etc.

Após todas as visitas e reuniões no Benjamin Constant, foi montado um cadastro com 45 usuários em potencial. Deste total, dois foram utilizados em testes-piloto e quinze foram selecionados para os testes propriamente ditos (veja tabela), de acordo com os critérios de idade e percepção visual mencionados anteriormente.

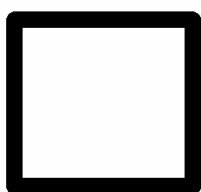
Todos os usuários selecionados, com exceção de 08, 09, 10, 12 e 14, têm pouca ou nenhuma experiência com computadores (veja coluna correspondente na tabela). Os que pouco tiveram contato geralmente fizeram apenas o curso introdutório de *Dosvox* (Dosvox, 1998) no próprio instituto e raramente voltaram a utilizar o programa novamente.

7.2 Definição dos Testes

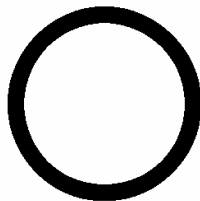
Os testes implementados foram divididos em três categorias principais: testes geométricos, testes espaciais e testes com texturas. Esses três tipos de testes serão detalhados a seguir.

7.2.1 Testes Geométricos

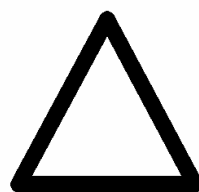
Os testes geométricos têm como principal objetivo avaliar se a interface permite ao usuário identificar a forma de objetos geométricos simples, tais como quadrado, círculo e triângulo (Figura 7.1). Como não se sabia a princípio se os usuários cegos desde o nascimento conheciam o nome de tais objetos, optou-se por apresentá-los da forma mais direta possível: foram recortadas três folhas de um material semelhante ao isopor (utilizadas principalmente para construção de maquetes) com tamanho de 25cm x 20cm e 0,50cm de espessura. Em cada folha foi esculpido em baixo relevo o contorno de um quadrado, círculo e triângulo. Desse modo, mesmo aquelas pessoas que não têm noções de geometria poderiam sentir os objetos no isopor.



(a)



(b)



(c)

Figura 7.1: Testes Geométricos

Os testes geométricos consistem em entregar os objetos de isopor ao usuário, pedir para ele senti-los e, a seguir, apresentar os mesmos objetos representados no computador. Após o usuário sentir tal objeto, pedia-se para que ele mostrasse o isopor correspondente. A idéia inicial era apresentar a sequência quadrado, círculo e triângulo no computador.

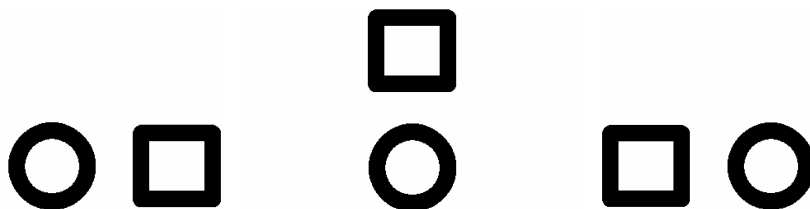
As hipóteses consideradas neste teste são:

- **HIPÓTESE H1.** A força aplicada pelo dispositivo reativo será capaz de manter o usuário próximo da borda do objeto;
- **HIPÓTESE H2.** O usuário será capaz de identificar a forma do objeto percorrendo seu contorno.

7.2.2 Testes Espaciais

Os testes espaciais têm como principal objetivo avaliar se a interface permite ao usuário identificar a posição de um objeto em relação a outro e ter noção de tamanho. Primeiro é colocado um círculo no centro da tela e um quadrado ao redor do círculo, próximo da borda da tela. O círculo sempre fica no meio e o quadrado pode estar à direita, à esquerda, acima e abaixo do círculo. Após mostrar cada uma das quatro situações, o usuário responde qual a posição do quadrado em relação ao círculo (Figura 7.2 a, b, c, d).

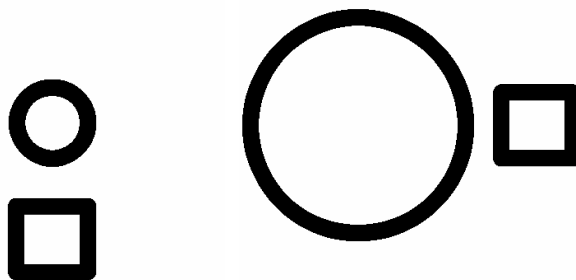
Em uma outra situação, é mostrado um círculo à esquerda e um quadrado à direita. O círculo foi propositalmente construído maior que o quadrado. Após a análise do usuário, pergunta-se qual dos dois objetos é o maior (Figura 7.2e).



(a) quadrado à direita

(b) quadrado acima

(c) quadrado à esquerda



(d) quadrado abaixo

(e) círculo maior que quadrado

Figura 7.2: Testes Espaciais

As hipóteses consideradas neste teste são:

- **HIPÓTESE H3.** O dispositivo reativo será capaz de transmitir ao usuário a noção de tamanho e posição relativa;
- **HIPÓTESE H4.** O usuário será capaz de movimentar com facilidade o *cursor* entre o círculo e o quadrado.

7.2.3 Testes com Texturas

Os testes com texturas têm o principal objetivo avaliar se a interface permite ao usuário diferenciar uma textura de outra. Conforme havia sido dito no Capítulo 6, o dispositivo háptico utilizado não é adequado para a simulação de texturas virtuais, por isto não houve a preocupação de representar virtualmente texturas existentes no mundo real. Se o usuário for capaz de diferenciar uma textura de outra, o objetivo foi atingido. Foram utilizadas nos testes as três texturas implementadas: *bump* (Figura 7.3a), *friction* (Figura 7.3b) e *vibration* (Figura 7.3c).

Os testes com texturas consistem em apresentar ao usuário uma tela totalmente preenchida com uma das três texturas. Após sua análise, esta tela é substituída por uma outra, dividida em quatro regiões (uma em cada canto) preenchida com as três texturas e mais uma textura vazia (Figura 7.3d). O usuário terá que indicar em qual região está a textura que analisou anteriormente. O mesmo processo se repete para as demais texturas.

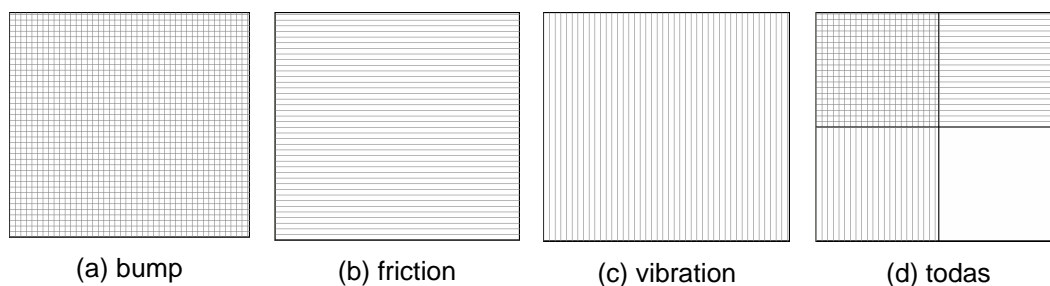


Figura 7.3: Testes com Texturas

Espera-se que o teste com texturas seja, pelo menos em tese, mais intuitivo para o usuário em comparação com os demais testes. Nos testes geométricos, por exemplo, a ausência do procedimento exploratório ótimo para a percepção da

forma global (*Enclosure*) faz com que os usuários utilizem o procedimento exploratório imediatamente inferior suportado pelo dispositivo háptico (*Contour Following*). Isto não acontece com o reconhecimento de texturas. Neste caso, o procedimento exploratório ótimo (*Lateral Motion*) é disponibilizado pelo dispositivo háptico, não sendo necessário, portanto, recorrer a um procedimento exploratório não-ótimo.

Tendo em vista esta característica tornar o reconhecimento de texturas uma tarefa potencialmente mais simples para o usuário, optou-se por avaliar neste teste não apenas a capacidade de diferenciação de texturas, mas também as limitações do dispositivo na representação de texturas virtuais. Por este motivo, foram escolhidas duas texturas semelhantes: *bump* e *vibration*. Espera-se que elas sejam facilmente confundidas, caso o usuário não perceba que a primeira só atua com o *cursor* em movimento. Esta semelhança será avaliada com mais detalhe nos testes.

As hipóteses consideradas neste teste são:

- **HIPÓTESE H5.** O usuário será capaz de identificar a textura *friction*;
- **HIPÓTESE H6.** O usuário não será capaz de diferenciar a textura *bump* da textura *vibration*.

7.3 Testes-Piloto

Inicialmente, os testes foram aplicados em duas pessoas com visão normal, porém com os olhos vendados. Essas pessoas possuíam bastante experiência com computadores como usuário, mas nunca utilizaram o *joystick* reativo anteriormente. Não houve preocupação em simular o ambiente e as condições onde os testes oficiais seriam realmente realizados. Também não era objetivo avaliar possíveis diferenças entre pessoas videntes com olhos vendados e deficientes visuais. Estes testes informais serviram apenas estimar a duração total da sessão de testes (o máximo esperado era de uma hora) e verificar se as instruções lidas sobre os testes estariam claras para os usuários.

O tempo total gasto por cada usuário foi de aproximadamente 50 minutos, incluindo um breve tutorial com o programa, os testes propriamente ditos e uma rápida entrevista de avaliação (esses tópicos serão descritos a seguir). Os usuários relataram que não ficaram cansados após o teste e entenderam perfeitamente as instruções dadas pelo avaliador. Por isso, até esse momento, nenhuma alteração

foi introduzida nos testes, pois não foram encontrados indícios de justificassem mudanças.

A seguir, foi realizado um primeiro teste-piloto com uma deficiente visual (totalmente cega) do Instituto Benjamin Constant. Neste teste houve a preocupação de utilizar o mesmo ambiente e condições para os testes oficiais. Toda a sessão foi gravada em vídeo para posterior análise. O programa de teste também gerou um arquivo com o registro de diversas informações quantitativas a respeito do teste, conforme será descrito mais à frente. Este primeiro teste-piloto foi essencial para acertar alguns detalhes até então não observados.

Inicialmente pensou-se em gravar as sessões de teste em vídeo digital (DV) pois a observação da fisionomia, o comportamento e as reações dos usuários poderiam ser relevantes para análise. Entretanto, devido ao grande volume de gravações e as dificuldades para localizar e selecionar os trechos importantes, este método foi logo abandonado. Além disso, ao saber que seu rosto estava sendo filmado, a usuária ficou visivelmente inibida. Ao invés de utilizar gravação em vídeo, optou-se por registrar apenas o que acontece na janela do programa de teste e a voz do usuário através de um programa de captura de telas, chamado *SnagIt* (TechSmith, 2003). Este programa gera um arquivo de animação (AVI) que é facilmente visualizado no computador.

Outra questão relevante mostrada no primeiro teste-piloto foi a obviedade dos testes geométricos. Conforme descrito anteriormente, neste teste o usuário recebe três folhas de isopor com objetos esculpidos em baixo relevo. A seguir o computador apresenta os mesmos objetos e pede para o usuário dizer a qual folha de isopor ele corresponde. O primeiro objeto é um quadrado, o segundo um círculo e o último um triângulo. Notou-se que o triângulo foi muito rapidamente identificado, pois após o usuário reconhecer os dois primeiros, o último só poderia ser o triângulo, pois foi única folha de isopor que sobrou. O teste foi modificado para incluir a repetição do quadrado. A ordem foi alterada para: quadrado, círculo, quadrado e triângulo. Obviamente, o usuário foi avisado que poderia haver repetições e algum objeto poderia até nem ser apresentado.

Nos testes espaciais, notou-se uma repetição desnecessária de tarefas essencialmente semelhantes: pedia-se para o usuário identificar a posição de um quadrado em relação a um círculo. As possíveis posições seriam quadrado à direita, à esquerda, acima ou abaixo do círculo (essa informação era dita ao

usuário). A ordem apresentada era quadrado à direita, acima, à esquerda e abaixo. Optou-se por substituir a última situação (quadrado abaixo do círculo) pelo quadrado à direita e abaixo do círculo. Nesse caso, seria dito ao usuário que o quadrado além de poder estar à direita, à esquerda, acima ou abaixo do círculo, também poderia estar em qualquer um dos outros cantos.

Nos mesmos testes espaciais, notou-se também uma ligeira confusão com relação aos termos acima e abaixo. Para os que enxergam, parece ser mais natural associar o movimento do *joystick* para frente, ou seja, no plano horizontal, ao movimento do *cursor* para cima, ou seja, no plano vertical. Na tela do computador é exatamente isso que acontece. No entanto, especialmente para aqueles que nunca enxergaram, não faz muito sentido tal analogia, pois eles não sabem o que acontece na tela do computador. Portanto, para evitar possíveis embaraços, como este que ocorreu no primeiro teste-piloto, ao dar qualquer instrução ao usuário, optou-se por substituir os termos “acima” por “à frente” e “abaixo” por “atrás”. Desta forma, não há necessidade de o usuário realizar o mapeamento de sistema de coordenadas.

Como foram feitas mudanças significativas nos testes, optou-se por realizar mais um teste-piloto, apenas para confirmar se as alterações foram satisfatórias. Novamente, este teste foi realizado no Instituto Benjamin Constant com um deficiente visual (também totalmente cego) e sob as mesmas condições em que seriam realizados os testes oficiais. Por se tratar apenas de uma validação, optou-se por realizar este teste na véspera dos testes definitivos.

Os resultados obtidos foram satisfatórios, especialmente a substituição da filmagem pelo programa de captura de tela. A única mudança realizada após o segundo teste-piloto foi na parte inicial da sessão de testes. Até o momento, ao receber o usuário na sala de testes, o avaliador lia um texto padrão (veja Apêndice A) de boas vindas e explicações sobre o que iria acontecer durante a sessão. Esta parte tem a duração de aproximadamente três minutos. Mas, como esse mesmo texto seria repetido quinze vezes, optou-se por gravar a voz do avaliador e apresentar esta gravação. Apesar deste procedimento tornar o início da sessão de teste um pouco impessoal, não iria interferir nos testes propriamente ditos, pois logo a seguir o avaliador estaria constantemente interagindo com o usuário.

7.4 Aplicação dos Testes

Os testes foram realizados em uma sala reservada do Instituto Benjamin Constant. Nesta sala ficavam apenas o usuário e o avaliador sentados ao redor de uma mesa (Figura 7.4). O *notebook* ficava sobre a mesa e com a tela virada para o avaliador. O usuário sentava-se logo ao lado e o *joystick* ficava próximo a seu corpo, podendo ser facilmente manipulado como braço na horizontal.

Uma câmera de vídeo foi colocada atrás do avaliador enquadrando quase todo o corpo do usuário. O uso deste equipamento limitou-se apenas a registrar alguns trechos das sessões (especialmente o início) para que posteriormente um vídeo pudesse ser produzido. Entretanto, antes de começar cada sessão, a avaliador pedia o consentimento do usuário para fazer a filmagem e também o avisava quando isso iria acontecer.

Um microfone de lapela também foi preso em sua roupa. Para que os usuários não ficassem esperando, foi marcado um horário dedicado para cada um. Cada sessão durou um pouco menos de uma hora.

A sessão de teste foi dividida em três etapas distintas. Inicialmente, ao receber o usuário na sala de testes, o mesmo era conduzido à mesa, sentando-se em sua cadeira apropriada. A seguir, o usuário ouvia uma gravação contendo instruções básicas a respeito da sessão de testes (ver Apêndice A). Esta gravação tem a duração de três minutos.

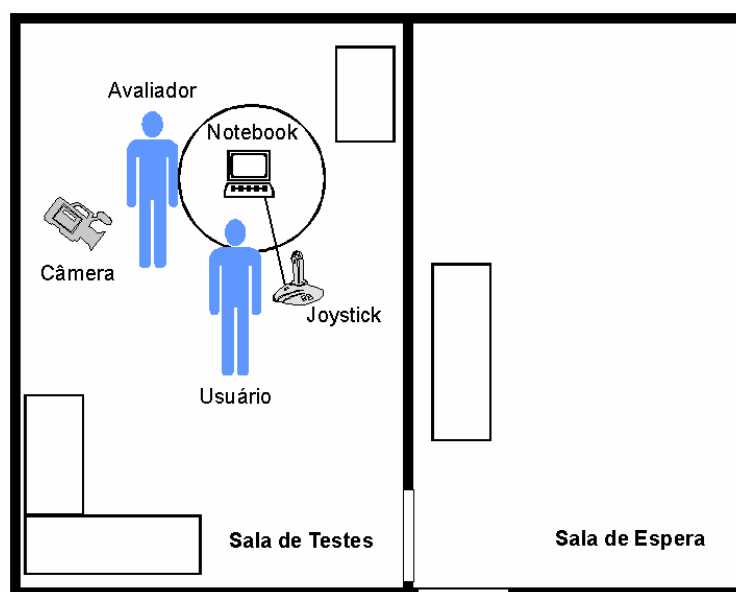


Figura 7.4: Ambiente de Testes

Depois da mensagem introdutória, o usuário tem a oportunidade de experimentar os conceitos mais importantes da interface. O avaliador conduz o usuário na realização de exercícios simples, tais como sentir uma linha horizontal, vertical, inclinada e *polyline*. Depois o avaliador apresenta vários objetos no computador e mostra como mover entre os objetos. Finalmente o usuário experimenta texturas.

Esses exercícios têm a finalidade de apresentar de uma maneira bem simples como utilizar o programa e o *joystick* reativo. Esse pode ser considerado o primeiro contato efetivo do usuário com todo o aparato computacional envolvido (alguns usuários tiveram a oportunidade de experimentar rapidamente, mas de modo desordenado, o protótipo do projeto durante a etapa de seleção de usuários em potencial).

Durante os exercícios, os usuários podiam perguntar, tirar dúvidas, experimentar novamente etc. e eram sempre assistidos e conduzidos pelo avaliador. A duração desta etapa variou de 6 a 20 minutos, dependendo da facilidade do usuário em entender os conceitos apresentados.

A seguir, os testes propriamente ditos eram aplicados ao usuário. A partir deste momento, o avaliador deixava de orientar, conduzir e responder perguntas do usuário. Qualquer interpretação era deixada a seu cargo. Entretanto, o avaliador procurava deixar o usuário bem à vontade, sem controlar estritamente o tempo do teste ou cobrar maior desempenho. Quando o usuário demonstrava impaciência ou dificuldade em realizar alguma tarefa, o avaliador procurava motivá-lo. No entanto, mesmo após muito tempo de insistência, alguns usuários não conseguiam avançar em determinadas situações,. Neste caso, para não prolongar demais a sessão, foi necessária a intervenção do avaliador no sentido de passar para o teste seguinte.

Conforme discutido anteriormente, após a avaliação dos testes-piloto, os testes definitivos foram organizados na seguinte ordem:

1. *Teste geométrico #1 (G1)*: Identificar a forma de um quadrado;
2. *Teste geométrico #2 (G2)*: Identificar a forma de um círculo;
3. *Teste geométrico #3 (G3)*: Identificar a forma de um quadrado (idêntico ao teste G1);
4. *Teste geométrico #4 (G4)*: Identificar a forma de um triângulo;

5. *Teste espacial #1 (S1)*: Identificar a posição de um quadrado à direita de um círculo;
6. *Teste espacial #2 (S2)*: Identificar a posição de um quadrado à frente de um círculo;
7. *Teste espacial #3 (S3)*: Identificar a posição de um quadrado à esquerda de um círculo;
8. *Teste espacial #4 (S4)*: Identificar a posição de um quadrado à direita e atrás de um círculo;
9. *Teste espacial #5 (S5)*: Identificar o objeto de maior tamanho (círculo ou quadrado);
10. *Teste com texturas #1 (T1)*: Identificar a textura *bump*;
11. *Teste com texturas #2 (T2)*: Identificar a textura *friction*;
12. *Teste com texturas #3 (T3)*: Identificar a textura *vibrating*.

A duração dos doze testes propostos acima variou de 5 a 40 minutos, de acordo com a habilidade do usuário. Durante todo esse momento, a tela do *notebook* e o áudio (voz do usuário e avaliador) eram registrados em um arquivo AVI e optou-se por registrar alguns trechos curtos das sessões em DV.

Finalmente, após o término dos testes propriamente ditos, o avaliador fazia uma rápida entrevista com o usuário. O questionário utilizado está apresentado no Apêndice B. Este questionário inclui basicamente perguntas objetivas e algumas poucas perguntas subjetivas. Além deste modelo padrão, o avaliador aproveitou este momento para questionar algum fato relevante que por ventura tivesse ocorrido durante a sessão de testes.

A entrevista também foi gravada (formato MP3) e sua duração variou de 4 a 48 minutos, de acordo com o grau de participação e interesse do usuário. No total, a sessão completa de teste (prática inicial, testes propriamente ditos e entrevista de avaliação) teve duração sempre inferior a uma hora. Em apenas duas situações, devido ao grande interesse do usuário e profundidade da discussão, a sessão durou mais de uma hora.

Por questões de ética e respeito à privacidade dos usuários, todos foram avisados, no início da sessão, que os testes seriam gravados em áudio e alguns trechos em vídeo. Entretanto, nenhuma informação pessoal seria divulgada (ver Apêndice A). Durante a sessão, os usuários eram sempre avisados pelo avaliador quando a câmera era ligada. Além disso, após a sessão, o avaliador solicitava aos usuários permissão para utilizar trechos gravados na produção de um vídeo. Nenhuma objeção foi feita pelos quinze usuários que participaram dos testes finais.