

5 Métodos e Técnicas de Pesquisa

5.1 Delineamento da Pesquisa

5.1.1 Tema

Usabilidade nas tarefas de busca, aquisição e avaliação nas lojas de aplicativos para dispositivos móveis

5.1.2 Problema

No capítulo anterior foi demonstrado como o desenvolvimento de aplicativos e soluções voltadas para dispositivos móveis possuem particularidades em relação ao desenvolvimento voltado para *desktop* (item 4.2). Algumas dessas particularidades são inerentes ao aspecto móvel desses dispositivos, outras são relacionadas as soluções desenvolvidas para aqueles dispositivos, tal como hardware e software do aparelho. Dessa forma, pode-se observar que a forma de navegação e interação nas lojas de dispositivos móveis para as tarefas de busca, aquisição e avaliação dos aplicativos pode variar de loja para loja.

Caso as lojas não sigam as recomendações de usabilidade, os usuários podem encontrar dificuldades em escolher, adquirir e avaliar aplicativos nas lojas de dispositivos móveis, frustrando-se durante essas tarefas. Conforme apontado anteriormente, essas dificuldades ocorrem quando os usuários possuem uma expectativa que não é cumprida pelos sistemas, seja por uma dissonância na comunicação a respeito do serviço ou porque os usuários não conseguem desenvolver os modelos mentais esperado pelos desenvolvedores do sistema.

5.1.3 Objeto da Pesquisa

O objeto da pesquisa é a interação entre usuários e os sistemas das lojas de aplicativos para dispositivo móveis.

5.1.4 Hipóteses e Variáveis

A hipótese considerada na pesquisa é a de que, se não forem consideradas as particularidades, potencialidades e limitações do usuário e do contexto na interação, em especial nas tarefas de busca, aquisição e avaliação, podem ocorrer erros, confusões e não atingimento dos objetivos da interação que por sua vez resultam em insatisfação e frustrações no uso da loja.

Para verificação dessa hipótese, tem-se como variável independente as lojas de aplicativos para *smartphones* desenvolvidas pelos fabricantes desses dispositivos ou do sistema operacional neles presentes. Essas variáveis não são influenciadas ou alteradas por nenhum dos procedimentos da investigação, nem por qualquer ação por parte dos participantes da pesquisa.

Já as variáveis dependentes são **os erros, confusões, não atingimentos dos objetivos** assim como as **insatisfações e frustrações** dos usuários que sejam decorrentes das diferentes interações, ou seja, da realização das tarefas de busca, aquisição e avaliação nas lojas, que é a variável independente. Através de uma relação causal, tenta-se identificar os elementos e situações que possam indicar as respostas para as questões de pesquisa e conseqüentemente refutar ou confirmar a hipótese levantada.

5.1.5 Objetivos

Objetivo geral consiste em propor um conjunto de parâmetros e recomendações de usabilidade para os processos de busca, aquisição e avaliação dos aplicativos nas lojas de aplicativos para dispositivos móveis a fim de melhorar a Experiência do Usuário (UX). Para isso, é necessário também cumprir os objetivos específicos, tais como:

- Analisar os modelos de distribuição de conteúdo dos principais competidores atuais: Loja *Nokia*, *iPhone App Store*, *Google Play Store* e *BlackBerry App World*.
- Identificar o perfil dos usuários de lojas de aplicativos de *smartphones* no Brasil.
- Identificar as preferências, opiniões e satisfação desses usuários em relação as lojas de aplicativos.
- Avaliar o processo de busca, aquisição e avaliação de aplicativos nas lojas de aplicativos.
- Identificar e relacionar os pontos positivos e negativos de cada loja nos processos de busca, aquisição e avaliação com base em técnicas e critérios de Ergonomia e Usabilidade.

Tem-se como objetivos operacionais a aplicação de métodos e técnicas de pesquisa, quantitativas e qualitativas, que levantem dados sobre a opinião, atitude e satisfação dos usuários de *smartphones* em relação às lojas de aplicativos oficiais disponíveis para estes. Para isso, foram utilizados:

- Pesquisa quantitativa com questionário *on-line* composto de perguntas abertas e fechadas
- Teste de Usabilidade
- Análise dos resultados dos testes de Usabilidade e dos questionários e a partir destes, desenvolvimento de recomendações para as lojas de aplicativos nas atividade de busca, aquisição e avaliação de aplicativos.

5.1.6 Justificativa e aplicabilidade

Apesar de várias pesquisas e resultados para melhoria de *Usabilidade na área de e-commerce*, ainda não há estudos voltados para as lojas de aplicativos móveis a fim de se definir as semelhanças e singularidades destas em relação às lojas de *e-commerce*. Esse estudo objetiva suprir essa lacuna tendo em vista as possibilidades e limitações desse meio. As conclusões e os produtos da pesquisa servirão como:

- Referência para decisões, abordagens e escolhas projetuais nos processos de desenvolvimento de aplicativos para dispositivos móveis

que tenham foco na distribuição via lojas de aplicativos, cujo objetivo seja a melhoria da UX.

- Contribuição metodológica para a área de Ergonomia e Interação Humano-Computador no que se refere a avaliação de sistemas em dispositivos móveis
- Referência para a definição de parâmetros que sirvam para a melhoria da experiência dos usuários das lojas de aplicativos móveis para operadoras, fabricantes e desenvolvedores.
- Contribuição com a identificação dos problemas particulares do usuário brasileiro, podendo estes serem usados como argumentos para melhoria deste serviço por agências reguladoras e órgãos de proteção do consumidor frente aos prestadores desse serviço.

5.2

Método, Técnicas e Procedimentos

O método hipotético-dedutivo se inicia com a problematização de um determinado fenômeno que fora anteriormente observado pelo pesquisador. Após essa problematização, uma hipótese é levantada e então confrontada com os dados empíricos daquele fenômeno com vias de confirmar, ajustar ou refutar a hipótese levantada (CIRIBELLI, 2000).

A pesquisa pretende alcançar seu objetivo através da análise de um fenômeno específico, a interação nas lojas de aplicativo. A hipótese formulada em relação a esse fenômeno precisa ser confirmada, ajustada ou refutada. Para isso, serão observadas as relações de uma determinada população, os usuários de lojas de aplicativos, com um objeto específico, a interação com as lojas. Essas relações são as variáveis incontroladas e o objeto a variável controlada. Os achados dessa relação podem ser considerados como os dados empíricos que vão confirmar, ajustar ou refutar a hipótese.

5.2.1

Questionário

Para realização de uma pesquisa de interação humano-computador, um dos primeiros passos é definir o perfil do usuário (MAYHEW, 1999). Uma das maneiras de

se fazer isso é através de uma pesquisa de campo ou pela análise de dados previamente levantados. Embora haja alguma informação a respeito dos usuários de *smartphones* no Brasil, uma pesquisa de campo mais detalhada se faz necessária pois nem todas as questões abordadas por essa pesquisa podem ser respondidas com as informações atualmente disponíveis.

Esta pesquisa de campo procurou identificar as informações sócio-demográficas, a opinião geral e a satisfação dos usuários brasileiros quanto às lojas de aplicativos disponíveis para seus *smartphones*. Era importante saber também, se esse tipo de informação varia de fabricante para fabricante e em qual medida. Portanto, um ponto extremamente relevante foi escolher uma ou mais técnicas que possibilitavam o levantamento das informações necessárias da melhor forma possível, considerando-se para isso o nível de especificidade da informação, o tempo e os recursos disponíveis para realização de tal atividade.

Para tal, foi escolhida a técnica de questionário em uma de suas especificidades que é a sua aplicação online, via internet. Essa é uma técnica tradicional de pesquisa que consiste em apresentar uma série de perguntas a um grupo específico de pessoas, que para fins de pesquisa, é chamado de população. Um questionário pode ser aplicado de modo físico, ou de modo virtual, e seu objetivo é medir e coletar a opinião dessa população a respeito de um tópico específico (LAKATOS & MARCONI *apud* ABREU, 2009; GIL, 1985; MAYHEW, 1999; OPPENHEIM, 1992).

5.2.1.1 Propostas e Objetivos

O objetivo do questionário era identificar o perfil dos usuários das lojas de aplicativos para dispositivos móveis assim como suas atitudes, opiniões e comportamentos em relação as lojas *iPhone App Store*, *Google Play Store*, *Loja Nokia* e *BlackBerry App World*.

Para tanto, as perguntas desenvolvidas buscaram primeiramente identificar se o respondente possuía um *smartphone* e se era usuário de lojas de aplicativos. Confirmando-se isso, tentava-se delinear como esse usuário usa a loja de aplicativos, as frequências e seus tipos de interação. Além disso, outras perguntas averiguavam se o respondente entendia o modelo de distribuição de conteúdo e quais funcionalidades e processos, caso mais de um fosse possível, são mais usados nesse ambiente. Buscou-

se também discernir as principais tarefas, falhas e omissões desses sistemas, no que concerne a interação do usuário com o sistema.

5.2.1.2 **Questões da Pesquisa**

Como o objetivo do questionário era identificar o perfil do usuário brasileiro das lojas de aplicativos para *smartphones*, as perguntas primeiramente filtravam usuários e não-usuários dessas lojas. As questões eram mais voltadas para os usuários de lojas e tentavam identificar:

- Qual a frequência de uso da loja de aplicativo?
- Qual ou quais processos eram mais comuns nas atividades de busca, aquisição e avaliação dos aplicativos?
- Os processos de decisão e busca de informação se limitavam as informações disponíveis no ambiente das lojas ou o usuário fazia uso de outras fontes de informação? Caso positivo, quais fontes?
- Qual era a satisfação, atitude e opinião do usuários com a loja do seu *smartphone*?
- Qual era o perfil sócio-demográfico do usuário das lojas de aplicativos?

5.2.1.3 **Estrutura do Questionário**

O questionário foi dividido em quatro módulos principais (figura 5.1). O primeiro módulo consistia em apenas uma pergunta, mas nele se encontravam várias informações a respeito do questionário, tais como o objetivo da pesquisa, que instituição a realizava, o tempo para conclusão do questionário (obtido com base no tempo gasto pelos usuários do teste piloto) e para qual público se direcionada o questionário. Essas são algumas das informações que ajudam o respondente a ter confiança na pesquisa e no questionário, ou desistir logo caso o tema não seja de seu interesse (GIL, 1985; MAYHEW, 1999; OPPENHEIM, 1992).

Dependendo da resposta da pergunta desse primeiro módulo, o usuário era direcionado para uma das versões do módulo 2. Havia quatro versões desse módulo, composto das mesmas perguntas, mas cada versão era dedicada a uma loja em particular e fazia uso dos termos particulares de cada loja. Se o usuário respondesse

Apple na primeira pergunta, ele era direcionado para o módulo da *iPhone App Store*, se respondesse *Nokia*, seria o da Loja *Nokia*.

No segundo módulo verificava-se se o usuários conhecia e utilizava a loja de aplicativos oficial de seu *smartphone*. Caso isso se confirmasse, o respondente era direcionado para o terceiro módulo, que continha as questões de uso, preferências e atitudes em relação às lojas. Ao terminar o terceiro módulo, seguia-se o quarto e último módulo com perguntas sócio-demográficas. Caso o usuário respondesse no segundo módulo que ele não fazia uso da loja de aplicativos, ele era direcionado do segundo para o quarto módulo. Ao terminar de responder o quarto módulo, uma mensagem de agradecimento era mostrada para o usuário indicando também o fim do questionário. As telas do questionário e a lista de perguntas encontram-se no apêndice.



Figura 5.1 – Estrutura do Questionário

5.2.1.4 Teste Piloto

Conforme sugerem Gil (1985), Mayhew (1999) e Oppeheim (1992), o questionário foi revisado pelos pesquisadores afim de se verificar se este cumpria os objetivos a que se propunha. Após essa verificação, os mesmo autores aconselham a realização de um teste piloto. Para essa etapa foram selecionados nove usuários que se encaixavam no perfil da pesquisa.

Mesmo tendo-se em mente que na versão final o questionário seria respondido pelos participantes sem a presença de um pesquisador, os testes pilotos foram conduzidos com a presença de um. Esse procedimento foi adotado pois assim se tornava possível a aplicação da técnica de *think aloud*, que consiste em solicitar que o participante "pense em voz alta", expondo verbalmente suas primeiras opiniões a respeito das perguntas apresentadas e a forma como são entendidas. O teste piloto presencial permitiu também que as expressões faciais e corporais dos participantes fossem observadas, o que ajudou a avaliar compreensão e eficácia do mesmo.

O teste piloto foi conduzido usando-se uma versão do questionário disponível no *site Wufoo*¹⁴. Embora aquela ainda não fosse a versão final, esta já apresentava a estrutura de perguntas, visual, textos introdutórios e de agradecimentos em suas versões mais avançadas. Estes só seriam mudados caso apresentassem alguma dificuldade para o usuário durante o teste piloto.

Nenhuma discrepância foi identificada durante o teste piloto. As reclamações e sugestões dos usuários diziam respeito apenas as terminologias utilizadas em algumas perguntas. Por mais que tenham sido usadas aquelas que tecnicamente correspondem as atividades e funções realizadas nas lojas de aplicativos, optou-se, na maioria dos casos, por usar os termos que eram mais familiares aos usuários, mesmo estes sendo imprecisos. Segundo Oppenheim (1992), o questionário deve fazer uso de uma linguagem clara e familiar para o respondente, para que assim esse possa ser entendido como interessante, inteligente e amigável.

¹⁴ <http://www.wufoo.com>

5.2.1.5 Formato, Aplicação e Promoção

A forma de preenchimento do questionário é um ponto crítico quando se faz uso dessa técnica de pesquisa, pois o preenchimento incorreto pode acarretar a perda daquela entrada ou mesmo a tabulação de dados incongruentes com a opinião e atitude do respondente (MAYHEW, 1999; OPPENHEIM, 1992). Por isso é importante sempre deixar claro a forma como o questionário deve ser preenchido. Uma das razões da escolha do *site Wuffo* frente às outras possibilidades foi o fato de que este usava elementos de navegação e interação tradicionais da web, tais como tipografia, botões, *checkboxes*, *links*, etc. Com isso as instruções sobre como responder o questionário não precisaram ser por demais extensas. Ainda assim, alguns cuidados foram mantidos, tais como a indicação, após algumas perguntas, quando fosse possível marcar mais de uma resposta.

Os convites feitos e a promoção do questionário foram basicamente *online*. Para isso, fez-se uso de convites diretos por email para as lista de contatos assim como listas de discussão e grupos *online*. Afim de não gerar uma má impressão do questionário, escolheram-se as listas e grupos que tivessem alguma afinidade com o tema da pesquisa, como a Lista de Arquitetura de Informação e o grupo de discussão de *Android* no *Google Groups*.

Além disso, recorreu-se as três principais redes sociais em uso no Brasil, no momento da pesquisa. No caso, Facebook¹⁵, Twitter¹⁶ e Orkut¹⁷. Assim como no caso das listas, procurou-se grupos e comunidades que tivessem alguma relação com o tema da pesquisa.

Durante essa promoção, foram usadas algumas técnicas sugeridas na literatura especializada com vias do aumento da visibilidade do questionário¹⁸. Um ponto principal é o convite, que deve ser claro e explicar os pontos da pesquisa, as instituições envolvidas, o tempo médio para preenchimento do questionário e também a confidencialidade e anonimidade dos resultados. Para uma sinergia e por uma maior consistência, usou-se o mesmo nome no convite, título e barra de endereço do questionário. Com isso, procurava-se ressaltar a seriedade da pesquisa e deixar claro

¹⁵ <http://www.facebook.com>

¹⁶ <http://www.twitter.com>

¹⁷ <http://www.orkut.com>

¹⁸ Se o título do questionário era “Questionário de uso do Android Market” o endereço do questionários era <http://wilsonprata.wufoo.com/androidmarket>. O mesmo foi feito para o sistema da Apple, Nokia e BalckBerry.

para os respondentes que os convites não se tratavam de “spam”, tal como sugerem Gil, (1985); Mayhew (1999) e Oppenheim (1992).

5.2.1.6 Análise dos dados

O questionário ficou disponível para ser respondido do período de julho a setembro de 2012. Após a data determinada para fim do questionário, o mesmo foi removido do site. Antes disso, foram usadas as próprias ferramentas do *site Wuffo* para exportar os resultados de cada loja em planilhas separadas no formato xls, compatível com o programa Microsoft Excel. As planilhas foram analisadas separadamente em um momento inicial, e depois consolidadas e analisadas em conjunto. As perguntas abertas foram separadas e analisadas pontualmente.

5.2.2 Teste de Usabilidade

O Teste de Usabilidade foi planejado seguindo a metodologia proposta por Rubin e Chisnell (2008), como foi apresentado no capítulo 4 (item 4.3). Esses autores argumentam que testes de Usabilidade podem ser realizados em qualquer momento do ciclo de desenvolvimento de um produto de Interação Humano-computador. O que vai variar em cada etapa do desenvolvimento é o tipo de questões de pesquisa a serem aplicadas, o nível de detalhamento do produto e a relação entre o tempo de projeto disponível e o esforço necessário para implementação de mudanças e melhorias.

Por ser um teste com dois sistemas diferentes que já estão lançados, o teste foi um **teste comparativo**. Testes comparativos podem ser feitos em qualquer momento do ciclo de desenvolvimento do produto. Eles são úteis para comprar duas ou mais propostas de design para a interface, ou comparar o manual atual contra uma nova proposta ou ainda para comparar a proposta com a de um concorrente.

5.2.2.1 Propostas e Objetivos

O objetivo do teste de Usabilidade é avaliar as tarefas de busca, aquisição e avaliação nas lojas de aplicativos para dispositivos móveis. Após identificada a popularidade de cada sistema, tanto conforme divulgado por institutos e empresas de

monitoramento (REIS, 2012. DISTMO, 2012) quanto pela pesquisa de campo realizada (questionário), decidiu-se concentrar os esforços de pesquisa nos dois principais sistemas, sendo estes: o *Android* e *iOS*, cujo as lojas são a *Google Play Store* e *iPhone App Store*, respectivamente. Limitando-se a essas tarefas, procurou-se identificar:

- Pontos críticos nas entradas de dados para realização das tarefas.
- Expectativa e opiniões em relação aos sistemas de rotulação, navegação, organização e busca.
- Medir e avaliar preferências em relação aos dois sistemas

5.2.2.2

Questões da Pesquisa

A partir dos problemas de pesquisa foram desenvolvidas questões de pesquisa que pudessem ser solucionadas através de um teste de Usabilidade. Desse modo, questionou-se se:

- Os usuários apresentam dificuldade nas tarefas de busca, aquisição e avaliação de aplicativos dentro das lojas no seus *smartphones*?
- A estrutura e organização das lojas são claras e coerentes para o usuário durante todo o processo de execução das tarefas de busca, aquisição e avaliação?
- Os termos e rótulos utilizados nas lojas são claros e coerentes para o usuário durante todo o processo de execução das tarefas de busca, aquisição e avaliação?
- Os recursos, métodos e processos de busca são claros e coerentes para o usuário durante todo o processo de execução das tarefas de busca, aquisição e avaliação?
- Os elementos de navegação utilizados pelas lojas são claros e coerentes para o usuário durante todo o processo de execução das tarefas de busca, aquisição e avaliação?
- É claro para o usuário o modelo de navegação das lojas durante as tarefas de busca, aquisição e avaliação?
- Quais das informações presentes na loja são mais relevantes no momento da escolha de um aplicativo?

A partir dessas questões, tarefas passíveis de serem executadas em ambos os sistemas foram definidas para as seções de teste. Essas tarefas referiam-se a ações

cotidianas em uma loja de aplicativo e poderiam ser realizadas de mais de uma maneira. Desse modo poder-se-ia analisar não apenas a eficiência e eficácia na execução das tarefas como também as preferências dos usuários frente as diversas opções de interação que os sistemas ofereciam.

5.2.2.3 Tarefas e Cenários

Conforme comentado anteriormente, o uso de cenários facilita a compreensão das tarefas a serem realizadas pelos usuário. Primeiramente foram desenvolvidas cinco tarefas e então foi pensado um cenário para cada uma delas. As tarefas foram:

- Baixar o aplicativo *Angry Birds*.
- Comentar “Ótimo passatempo” no aplicativo *Angry Birds*.
- Baixar um aplicativo da área de destaque.
- Baixar um aplicativo de clima.
- Baixar um navegador para internet.

E os cenários correspondentes a cada uma delas:

- Cenário 1: Seus amigos andam comentando bastante sobre um aplicativo, na verdade um jogo, chamado *Angry Birds*. Parece ser um jogo bem casual, para jogar por alguns minutos enquanto espera em uma fila por exemplo. Você ficou sabendo que existe uma versão gratuita do aplicativo e decide experimentar. Instale o aplicativo no seu *smartphone*.
- Cenário 2: Você baixou o aplicativo *Angry Birds* e gostou bastante. Ele se mostrou um ótimo passatempo. Como o aplicativo é gratuito, você pensou que poderia ser interessante ajudar na divulgação do jogo. Uma forma de fazer isso é dando uma boa avaliação e fazendo um comentário positivo a respeito. Agora que você usou o aplicativo *Angry Birds*, faça uma avaliação na loja e insira um comentário dizendo Ótimo passatempo!
- Cenário 3: Você ficou sabendo que existe uma área de destaque na loja de aplicativos do seu *smartphone*. Esses aplicativos são sugeridos pela própria loja, então deve haver um bom motivo para estarem nessa lista. Já que você ainda esta procurando por aplicativos para seu *smartphone*,

por que não verificar o que a loja anda sugerindo. Veja se há uma área de destaque e tente baixar um aplicativo gratuito (o primeiro disponível)

- Cenário 4: Você não gosta de surpresas, principalmente quando surge uma friagem ou uma chuva inesperada. A gente sabe que a meteorologia não é 100%, mas melhor com ela do que sem ela. Já que agora você tem um *smartphone* com plano de dados você pode consultar a previsão do tempo para poder se planejar melhor. Alguns sites de previsão do tempo não estão adaptados para o celular, melhor se tivesse um aplicativo. Tente baixar um aplicativo de clima para o seu *smartphone*.
- Cenário 5: Você usa bastante seu celular para navegar na internet. Entretanto você percebe que nem todas as funcionalidades que você normalmente usa em seu computador, estão disponíveis no navegador do seu celular. Pesquisando sobre o assunto você descobre que pode instalar outro navegador e quem sabe, ter acesso a mais funcionalidades. Veja se existe algum outro navegador disponível para seu *smartphone*.

5.2.2.4 Perfil dos Participantes

Foi desenvolvido um questionário de recrutamento online para a seleção dos participantes do teste de Usabilidade. Esse questionário era bem breve, possuía apenas 10 perguntas e buscava filtrar os usuários que poderiam participar do teste de Usabilidade, com base no perfil que foi definido a partir da análise dos resultados do primeiros questionário aplicado. O principal critério para seleção desses participantes era que fossem usuários de *smartphones* com ou sem experiência no uso de lojas de aplicativos, mas que não haviam utilizado anteriormente as lojas que estavam sendo avaliadas.

As outras perguntas ajudavam a complementar o perfil do respondentes. Estas perguntavam sobre o gênero, idade, escolaridade e também disponibilidade para participar do teste. Um total de 16 participantes foram selecionados para o testes. Possuíam entre 18 e 40 anos, divididos igualmente entre homens e mulheres. O nível de escolaridade ficava entre o ensino médio completo e o superior completo. A participação era voluntária.

5.2.2.5 Sistemas Avaliados

Os usuários utilizaram as duas lojas. A loja do Android, *Google Play Store*, e a loja da *Apple*, *iPhone App Store*. Os *smartphones* que utilizados possuíam configurações bem parecidas, sendo estes o *Galaxy Ace* e o *iPhone 3GS* respectivamente.



Figura 5.2: Samsung *Galaxy Ace* e *iPhone 3GS*. Fonte: gsmarena.com

Esses *smartphones* possuem capacidade de armazenamento, resolução de tela e capacidade de processamento bem parecidas. Ambos possuem tecnologia de toque de tela parecida, identificando mais de um toque por vez. O Samsung *Galaxy Ace* é 10% menor e mais fino, possui uma frequência de banda maior para conectividade (COMPARE CELLPHONES, 2012). Ainda assim, as pequenas diferenças presentes não deveriam interferir no uso das lojas de aplicativos, pois são aplicativos que não requerem tanto esforço de processamento e medidas quantitativas de desempenho da loja não serão avaliadas, tais como tempo de resposta e quantidade de dados carregados.

5.2.2.6 Ambiente e Equipamento do Teste

A maioria dos testes foram conduzidos em uma sala da PUC-Rio. Nesse local havia as facilidades de controle de entrada e saída de pessoas, evitando distrações e perturbações no teste; rede de internet sem fio, o que permitia o acesso a ambas as lojas de aplicativos com paridade de performance; condições de iluminação, conforto

térmico e acomodação apropriada. As instalações se aproximavam bastante do modelo mais simples do laboratório de Usabilidade proposta por Rubin e Chisnell (2008).

Além da sala na PUC-Rio, foram realizados testes em mais dois locais, duas residências. Em ambos, foram mantidas as características descritas anteriormente, principalmente no que se refere ao isolamento e imperturbabilidade da sessão, conexão de internet e conforto do participante. O registro não foi comprometido graças as características móveis de todo equipamento utilizado para os testes.

Comparando com o equipamento sugerido para um laboratório simples de Usabilidade conforme sugerido por Rubin e Chisnell (2008), os artefatos e aparatos tecnológicos utilizados se aproximavam mais do laboratório portátil, conforme descrito pelo mesmo autor. Esse é um modelo com boa relação custo *versus* benefício e, devido as particularidades do teste em dispositivos móveis, se mostrou a melhor solução.

Portanto, foi utilizado um laptop para registro de áudio e vídeo da interação do usuário com os sistemas. Para possibilitar esse registro, foi adquirida uma webcam e foi desenvolvido um suporte em acrílico para fixação desta e do celular em teste. Um programa gratuito era utilizado para o registro em vídeo (*Debut*) e as sessões de *debriefing* eram gravadas utilizando-se o gravador do *smartphone*. Os formulários, roteiros e questionários eram poucos e cabiam em uma pasta. Todo o equipamento, analógico e digital, cabia em uma mochila simples, mesmo contando carregadores, adaptadores e extensões.

5.2.2.7 Procedimentos das Sessões do Teste

Após a realização dos testes piloto, foi identificado que a sessão de teste durava em torno de 50 minutos. Estes estavam divididos em:

- Apresentação, explicação da metodologia, do termo de consentimento, aplicação do questionário e entrevista pós-questionário (10 minutos)
- Tarefas do primeiro sistema (10 a 15 minutos)
- *Debriefing* do primeiro sistema (05 a 10 minutos)
- Tarefas do segundo sistema (10 a 15 minutos)
- *Debriefing* do segundo sistema (05 a 10 minutos)

Para condução da sessão de teste foi desenvolvido e utilizado um roteiro para o moderador, que era o próprio pesquisador. O roteiro, e outros documentos desenvolvidos para o Teste de Usabilidade, encontram-se nos anexos. Esse roteiro

buscava garantir que as instruções fossem as mesmas para todos os participantes do teste, minimizando o ruído nas informações. A primeira etapa do teste era, após recepção do participante e apresentação do moderador, a leitura e aceitação do termo de consentimento de participação por parte do participante, que descrevia o envolvimento deste, como seria o registro da sessão e o uso que se faria das informações levantadas. Caso o participante estivesse de acordo, solicitava-se a esse que preenchesse um breve questionário sobre uso de *smartphones* e dispositivos móveis, bem parecido com o questionário de seleção para o teste.

Após o usuário preencher esse questionário, revia-se cada pergunta solicitando para o respondente especificar e detalhar um pouco mais cada resposta. Esse cuidado visava confirmar o grau de experiência do usuário com o uso ou não das lojas de aplicativos em *smartphones* e *tablets*, visto que vários desses últimos utilizam lojas muito parecidas com as versões para *smartphone*. Como existem *tablets* com sistema operacional *Android* e *iOS*, e as lojas para esses *tablets* possuem os mesmos paradigmas de interação que a versão para *smartphone*, caso o usuário tivesse experiência com algum desses artefatos ele não seria considerado na avaliação dos resultados, visto que essa experiência poderia interferir no seu desempenho.

Passava-se então para explicação da sessão e início do teste em si. Mostrava-se como seria gravado a interação através da *webcam* fixada no suporte de acrílico, enfatizava-se a necessidade de se manter o celular sobre o suporte, podendo ser deixado sobre a mesa ou permanecer mantido em uma das mãos, desde que sempre com o celular sobre o suporte. Explicava-se o processo de *think aloud*, reforçava-se que o que estava sendo testado era o sistema operacional, a independência da pesquisa em relação aos fabricantes e o tipo de auxílio e esclarecimento que poderia ser feito pelo moderador.

O primeiro sistema era então apresentado para o usuário e uma breve explicação sobre seu uso era feita. Basicamente solicitava-se que o usuário abrisse a loja e saísse algumas vezes e fizesse algumas interações básicas com o sistema. Desse modo ele se familiarizava com a sensibilidade de toque da tela, metáfora do sistema operacional e saberia como abrir novamente a loja caso a fechasse acidentalmente durante o teste.

Durante as sessões foi observado que, a breve utilização do sistema para familiarização do usuário com o sistema operacional antes do início da realização das tarefas fora suficiente para que o usuário compreendesse satisfatoriamente os

processos de navegação. Desse modo a maioria dos problemas encontrados estavam relacionados a interação com a loja e não com o não entendimento de determinado processo particular daquele sistema.

Após isso, era lido o primeiro cenário e este era posto a frente do usuário para caso ele tivesse necessidade de consultar novamente o cenário e a tarefa a ser feita. Quando o usuário completasse ou desistisse daquela tarefa, passava-se para a seguinte até finalizar as cinco tarefas. Afim de mitigar os efeitos da aprendizagem do uso das lojas, as tarefas foram aleatorizadas entre os usuários. Ou seja, estes fizeram as mesmas tarefas mas em ordens diferentes.

Dessa modo encerrava-se o teste com o primeiro sistema, seguia-se o preenchimento do questionário SUS e a entrevista de *debriefing*, a qual era gravada o áudio para posterior análise. O SUS (*System Usability Scale*) é uma escala Likert de cinco pontos, composta por dez sentenças. Ele tem se mostrado uma escala simples, robusta e confiável para a avaliação de sistemas. Seu objetivo é avaliar qual, entre dois sistemas, é o mais usável (BROOKE, 2012). Foi desenvolvido em 1986 por John Brooke. Desde então vem sendo aplicada em testes de produtos, softwares, celulares e até mesmo nas páginas amarelas, tendo se tornado uma referência em avaliação de sistemas (SAURO, 2012).

Sendo uma escala de avaliação, o respondente deve escolher o quão ele concorda ou discorda com cada sentença proposta. As sentenças presentes na escala, assim como na maioria das escalas Likert, visam exemplificar atitudes e expressões extremas em relação ao uso do sistema. As perguntas foram escolhidas a partir de um grupo de cinquenta, desta selecionou-se aquelas que melhor possibilitavam o posicionamento claro da opinião do respondente e que tivessem menor número de respostas ambíguas. O SUS é geralmente aplicado logo após o uso do sistema que está sendo avaliado, antes de qualquer entrevista ou evento pós-teste. Solicita-se aos respondentes a preencher a primeira resposta que vier a mente (BROOKE, 2012).

Para obtenção do número equivalente àquela entrada, deve-se primeiramente somar a pontuação de cada item, que pontuam de 0 a 4. As perguntas 1, 3, 5, 7 e 9 são pontuadas com base na escolha marcada na escala menos 1. Sendo o mínimo zero, no caso de “Discordo totalmente”, e no máximo 5, se marcado “Concordo totalmente”.

As perguntas 2, 4, 6, 8 e 10 são pontuadas com o número 5 menos a pontuação da escolha marcada na escala. Sendo o mínimo zero, no caso de Concordo totalmente,

e no máximo 4, se marcado “Discordo totalmente”. Soma-se todos os resultados e multiplica-se o resultado por 2.5 para obtenção do valor final do SUS.

Conforme lembra Sauro (2012), apesar do resultado do SUS ser um número de 0 a 100, isso não significa uma porcentagem de uso. Um resultado 74 pode, através da técnica de normalização do resultados, ser convertido em um rank percentual de 70%. Já um resultado de 68 fica por volta de 51%.

O *debriefing* era composto por perguntas pré-estabelecidas e por dificuldades e particularidades no uso do sistema avaliado que haviam surgido durante a sessão. Passava-se então para o teste do segundo sistema seguindo os mesmos processos descritos a partir da introdução dos sistemas até o *debriefing*. Os cenários eram apresentados na mesma ordem que foram apresentados no primeiro sistema e o *debriefing* tinha foco no sistema em uso. Alguns usuários prontamente comparavam o uso dos dois sistemas, caso comentários desse tipo não surgissem era solicitado que os usuários se posicionassem em relação aos dois sistemas.

Ao final do teste do segundo sistema, desligava-se o equipamento de gravação, agradecia-se a participação do usuário, um brinde simbólico era dado – um *pendrive* de 2Gb – e então o participante era acompanhado até a saída.

5.2.2.8 Dados Coletados

Tendo em vista o objetivo da pesquisa de estudar o processo de busca, aquisição e avaliação de aplicativos com base em técnicas e critérios de Ergonomia e Usabilidade e de se identificar e relacionar os pontos positivos e negativos de cada loja nesses processos, decidiu-se analisar a completude das tarefas, o grau de eficiência e a satisfação do usuário com cada um dos sistemas.

Quanto aos resultados do teste, primeiramente os problemas levantados foram identificados e submetidos a um processo de classificação. Barnum (2002) *apud* AGNER (2007), propõem que esse processo pode ser feito de duas formas, ou *top-down* ou *bottom-up*. No primeiro caso, as categorias são definidas antes dos resultados serem analisados, geralmente escolhe-se ou defini-se uma série de heurísticas para que os problemas sejam organizados dentro dessas categorias. Na segunda abordagem as categorias são definidas após a análise dos resultados, os problemas são primeiramente levantados e a partir dos tipos de problemas surgidos que estes são agrupados em categorias emergentes.

Nessa pesquisa foi realizada a abordagem *bottom-up* visto que os diferentes modos de se realizar as tarefas de busca por um aplicativo, incluindo nesse caso as várias possibilidades de navegação dentro da loja, resultam em diferentes problemas. Além disso, é mais importante observar que não há uma forma correta e estabelecida de busca por um aplicativo. Com essa abordagem procurou-se identificar os modelos de uso mais freqüentes por partes dos usuários e os problemas que são encontrados nesse processo.

Para identificação e análise dos dados qualitativos, além do SUS foi feita a análise das entrevistas de *debriefing* onde se identificou os principais comentários positivos e negativos. Essas entrevistas foram de grande auxílio para a compreensão da experiência do usuário com as lojas de aplicativos, pois várias tarefas em que os usuários não apresentavam dificuldade na realização, estes tinham várias críticas e considerações.

5.2.2.9 Teste Piloto

Conforme sugerido pela bibliografia (RUBIN & CHISNELL, 2008. MORAES & SANTA ROSA, 2008), antes das sessões definitivas do teste de Usabilidade foi feito um teste piloto para validação do modelo proposto. Este foi realizado com cinco participantes, dois do sexo masculino e três do gênero feminino, sendo quatro participantes com idades entre 22 e 28 anos e um com mais de 40 anos.

Os sistemas avaliados no teste piloto foram os mesmos utilizados nas sessões finais do teste, e não apresentaram nenhuma discrepância entre o piloto e as sessões finais. Ainda no teste piloto, os cenários e as tarefas se mostraram claros e de rápido entendimento por parte de todos os participantes. Quanto a completude das tarefas, nem todos conseguiram realizar todas as tarefas sendo os maiores problemas relacionados a avaliação de aplicativos e para adicionar comentários, o que se repetiu nas sessões finais.

A completude da maior parte das tarefas não foi visto como um problema metodológico pois os sistemas testados são submetidos a esse e outros tipos de avaliação constantemente. Observou-se um equilíbrio no grau de dificuldades das tarefas pois todas foram completadas por pelo menos um dos usuários, e todos os usuários apresentaram alguma dificuldade em algum momento. Uma análise parcial dos

resultados mostrou que as tarefas poderiam indicar respostas para as questões de pesquisa definidas.

Com a realização do teste piloto foi possível identificar o tempo médio de duração da sessão, que ficou em volta de 50 minutos. Essa informação fora exposta no convite do teste para evitar mal entendidos e como critério para que os candidatos pudessem avaliar previamente sua disponibilidade. Alguns pequenos ajustes puderam ser feitos nos documentos. Erros de concordância foram identificados e sentenças que não estavam claras foram reescritas.

5.3 Conclusão do capítulo

Quando uma pesquisa procura resolver problemas relacionados a objetos recentes ou relativamente novos, como é o caso das lojas de aplicativos para *smartphones*, é preciso atentar as singularidades desse novo meio. Chapanis (1959) lista alguns pontos a se ter em mente quando lidamos com essas situações, sendo estes:

- Não subestime a capacidade humana. Se quisermos pensar no homem como uma máquina, devemos levar em conta que ele é uma máquina não linear, que muda continuamente, que esta sujeita a perturbações e a ruídos. Uma máquina que pensa, que tem atitudes e emoções, que vai ludibriar cada esforço feito para tentar defini-la em aspectos funcionais.
- Esteja alerta para novas técnicas de pesquisa. Nenhum livro de metodologia vai conseguir suprir as contínuas demandas de pesquisa. O mais importante para um bom pesquisador é imaginação, adaptabilidade e inventividade.
- Estabeleça cuidadosamente as condições de observação. Uma regra geral para todo experimento é anotar, e registrar, o máximo de informações possíveis.
- Mantenham uma atitude objetiva quando estudando as pessoas. Observe os erros com atenção e especialmente as informações que vão contra as ideias estabelecidas e que possam revelar falhas metodológicas. É preciso ser cético, observar e registrar os fatos primeiro para então analisá-los depois.
- Operacionalize o estudo do comportamento. A regra geral é de que podemos estudar as ações humanas apenas através de seu comportamento, não é

possível estudar a experiência do outro, visto que esta é única, individual e intransferível.

Os pontos apresentados se relacionam diretamente com a pesquisa realizada. Podemos entender, fazendo uma relação ponto a ponto com o objeto e objetivo da pesquisa, que:

- Conforme apresentado anteriormente as pessoas estão diante de um novo paradigma de distribuição e consumo de conteúdo. Este agora é móvel e virtual. Não apenas isso, mas as próprias relações de tempo, espaço e consumo são redefinidas e definem as práticas e modelos de uso desses sistemas
- A compressão de espaço e tempo na sociedade atual auxilia na divulgação, distribuição e debate de novas metodologias para lidar com esses novos paradigmas. A emergência das abordagens com foco na UX é um desdobramento dessas novas relações. A experiência do usuário é uma forma de se estudar e desenvolver aquilo que a bibliografia do campo não dava conta de explicar ou explicava superficialmente. Assim como as tecnologias evoluem e se especializam, uma lógica semelhante ocorre com as ciências que estudam aquelas tecnologias.
- Teste de Usabilidade é uma técnica tradicional nas pesquisas de Interação Humano-Computador, entretanto necessita de alguns ajustes para poder ser satisfatoriamente operacionalizada quando aplicada ao estudo de dispositivos móveis. Esse é um ponto claramente discutido na definição metodológica dessa pesquisa, onde além de explicadas foram justificadas as alterações necessárias.
- A definição de métodos tradicionais de avaliação busca minimizar os possíveis erros metodológicos. A vasta bibliografia de suporte aborda os problemas mais comuns e como minimizá-los. Para esgotar ainda mais essas possibilidades, testes pilotos foram realizados sempre que possível. Além disso, as etapas metodológicas que precisaram ser adaptadas receberam atenção redobrada.
- Chapanis (1959), apesar de não se referir diretamente ao problema, expõe claramente aqui porque é tão difícil de se estabelecer um método para avaliação do critério de satisfação, um dos três da Usabilidade. A experiência do usuário nasceu dessa dificuldade. Mesmo havendo um limite intransponível em determinado ponto, esse ponto precisa ser objetivamente esgotado em termos metodológicos. Os esforços de uma análise qualitativa com base nos comentários, atitudes e reflexões externadas pelos participantes dos testes,

tenta responder as questões de pesquisa sem perder de vistas as limitações de tal abordagem.

Interessante observar que Chapanis escreveu a respeito dos problemas da pesquisa com novas tecnologias onde o elemento humano era a principal variável, isso na década de 1950, e até hoje suas ponderações se mostram atuais. A partir dessa perspectiva entende-se que novas terminologias e métodos de pesquisa fazem parte da evolução do campo mas o mais importante deve ser sempre o elemento humano.

Na aplicação da metodologia definida, percebeu-se que esta era acima de tudo um processo para facilitar a compreensão das dificuldades dos usuários com os sistemas avaliados. O método escolhido levou em consideração o estado da arte da pesquisa de interação humano-computador, as possibilidades e limitações das tecnologias de comunicação móvel e as motivações, opiniões e atitudes do usuário em relação a essa tecnologia. Ainda que esteja clara a importância de se estudar a bibliografia consagrada pelo campo, as necessidades particulares de cada pesquisa exigem pequenas adaptações e melhorias para que o objetivo seja alcançado.

Os resultados encontrados após as mudanças realizadas nos métodos de pesquisa justificaram a necessidade das adaptações feitas e mostraram que a maioria das escolhas foram acertadas. As informações qualitativas expuseram a complexidade de se trabalhar uma pesquisa de interação humano-computador ao mesmo tempo que ilustraram a riqueza dos resultados e as muitas possibilidades de desdobramento e avanços nessa área. As informações quantitativas serviram como referência prática para os achados, visto que um objetivo específico havia sido fixado. Os próximos capítulos irão tratar os resultados encontrados através do que foi aqui definido, identificar os pontos mais importantes e os resultados mais relevantes para por fim confirmar ou refutar a hipótese levantada.