

3

Técnicas para avaliação de usabilidade

Este capítulo apresenta as técnicas para avaliações de usabilidade empíricas em produtos de mídia digitais desenvolvidos com a abordagem em cascata. Além disso, também apresenta as técnicas mais indicadas, pelos autores de IHC, para projetar e testar produtos desenvolvidos com métodos ágeis. Por fim, discute-se a validade das propostas para avaliação de usabilidade em produtos criados com a abordagem *agile*.

3.1.

Técnicas empíricas mais utilizadas na avaliação de usabilidade em produtos desenvolvidos com metodologia em cascata

Apresenta-se as técnicas para avaliações de usabilidade empíricas de produtos de mídia digitais: entrevista, questionário, grupo de foco, *card sorting*, análise da tarefa, avaliação cooperativa, teste de usabilidade e *eye tracking*. Considera-se essas técnicas como as mais utilizadas, tanto no mercado quanto na academia, para a realização de pesquisas em produtos desenvolvidos com a metodologia em cascata.

3.1.1.

Entrevista

Para MORAES e MONT'ALVÃO (2003), entrevista é uma técnica onde o investigador se apresenta frente ao investigado e lhe formula perguntas, com o objetivo de obter dados que interessam à investigação:

- Uma das partes busca coletar dados;
- A outra parte constitui-se em fonte de informação.

De acordo com MACHADO (2002) apud MIRANDA e MORAES (2005), entrevista é uma conversa entre duas pessoas, iniciada pelo entrevistador. Segundo RAMPAZZO (2002) apud MIRANDA e MORAES (2005), as principais vantagens da entrevista são:

- Pode ser utilizada com todos os segmentos da população, analfabetos ou alfabetizados;

- Há maior flexibilidade;
- Oferece maior oportunidade para avaliar atitudes e condutas;
- Oferece maior oportunidade para a obtenção de dados que não se encontram em fontes documentais e que sejam relevantes e significativos.

Segundo MORAES e MONT'ALVÃO (2003) a entrevista é o mais flexível de todos os métodos interrogativos de coleta de dados. Podem-se identificar os mais diversos tipos de entrevista. A classificação se faz a partir de critérios diversos, sendo que o mais usual se refere ao grau de estruturação. As entrevistas menos estruturadas desenvolvem-se de forma mais espontânea, sem que estejam sujeitas a um modelo apriorístico de interrogação. Já as entrevistas mais estruturadas são aquelas que predeterminam em menor ou maior grau as respostas a serem obtidas.

MORAES e MONT'ALVÃO (2003) ressaltam que, entre os vários tipos de entrevistas, distingue-se:

- **Entrevista não diretiva, ou entrevista aprofundada:** na qual a conversação é iniciada a partir de um tema geral sem estruturação do problema por parte do investigador;
- **Entrevista focalizada ou centrada:** na qual, dentro de hipóteses e de certos temas, o entrevistador deixa o entrevistado descrever livremente sua experiência pessoal a respeito do assunto investigado;
- **Entrevista semiestruturada:** aplicada a partir de um pequeno número de perguntas abertas;
- **Entrevista clínica:** eventualmente conduzida de maneira não-diretiva, cuja especialidade está na sua orientação em função do objetivo de interpretação sócio-psicológico da situação ou da personalidade dos sujeitos através de suas verbalizações.

Segundo JORDAN (1998), na técnica de entrevista, o pesquisador compila uma série de questões, propostas diretamente aos participantes, podendo utilizar uma entre as três categorias de entrevistas:

- **Entrevistas sem-estrutura:** pergunta-se uma série de questões, dando a oportunidade do respondente dirigir a discussão para os assuntos que considera mais importantes. Esta categoria é mais apropriada nas situações onde ainda não se conhece as questões de pesquisa a serem investigadas;

- **Entrevistas semiestruturadas:** o pesquisador possui uma ideia clara sobre as questões de pesquisa que são relevantes, assim como os assuntos que espera que os respondentes abordem. Durante o tempo todo, ele tenta assegurar que certos pontos serão discutidos. Mas, ao mesmo tempo, os respondentes também podem levantar questões que são importantes para eles;
- **Entrevistas estruturadas:** os respondentes devem escolher uma opção em um arranjo de respostas ou uma escala.

De acordo com GIL (1999) apud MIRANDA e MORAES (2005), a entrevista informal, dos tipos de entrevistas existentes, é o menos estruturado possível e só se distingue da simples conversação porque tem como objetivo básico a coleta de dados. O que se pretende com entrevistas deste tipo é a obtenção de uma visão geral do problema pesquisado, bem como a identificação de alguns aspectos da personalidade do entrevistado. A entrevista informal é recomendada nos estudos exploratórios, que visam abordar realidades pouco conhecidas pelo pesquisador, ou então oferecer visão aproximativa do problema pesquisado. Nos estudos desse tipo, com frequência, recorre-se a entrevistas informais com informantes-chaves, que podem ser especialistas no tema em estudo, líderes formais ou informais, personalidades destacadas, etc.

Não se pode esquecer que existe a possibilidade do entrevistado ser influenciado pelo investigador. MIRANDA (2005) relata que, ao iniciar as entrevistas, é importante que o investigador forneça uma série de informações para não intimidar o usuário, como:

- Quem é o investigador (cargo e função);
- O projeto para o qual serão usadas essas entrevistas;
- Os objetivos específicos dessas entrevistas;
- O que o investigador espera do usuário;
- Qual a duração das entrevistas;
- Como as informações dadas pelo usuário ajudarão no desenvolvimento do projeto.

Segundo MIRANDA (2005), ao conduzir as entrevistas, é possível utilizar o modelo “mestre (usuário) x aprendiz (investigador)”. Tal modelo implica em:

- A entrevista não deve ser muito estruturada;
- O mestre tem controle sobre a entrevista, mas permite que o aprendiz faça perguntas para melhorar seu entendimento;

- Como cada aprendiz tem sua própria maneira e ritmo de aprendizado, esse método funciona melhor com o usuário (mestre) interagindo com apenas um entrevistador (aprendiz).

MIRANDA (2005) também relata que durante as primeiras entrevistas:

- **Erro comum:** perder o foco, ou seja, deixar de lado o processo de extrair informações para começar a desenvolver ou a educar o usuário;
- **Deve-se evitar o erro:** para regular a expectativa do usuário, é preciso explicar previamente o propósito da entrevista e voltar a explicá-lo, caso necessário;
- **Quando o problema continuar:** no decorrer da entrevista, é possível oferecer a possibilidade de interromper o processo e continuar posteriormente em outro momento.

Além disso, deve-se aprender a linguagem e os jargões utilizados pelo usuário, e passar a usá-los assim que possível. Não se deve traduzir a linguagem do usuário para um linguajar técnico nem tentar que ele adote este linguajar. Ao final da entrevista, deve-se expressar satisfação e gratidão para os usuários.

3.1.2. Questionário

De acordo com OPPENHEIM (1992), questionar pessoas significa mais do que apenas fisgar um peixe particular, sendo necessárias diferentes tipos de iscas para profundidades distintas e sem saber o que está acontecendo logo abaixo da superfície. Portanto, OPPENHEIM (1992) afirma que um questionário não é um tipo oficial de formulário, nem tampouco é uma lista de perguntas reunidas sem um planejamento prévio. REA e PARKER (2000) definem questionário como uma série de perguntas, sem viés e bem estruturadas, que irão obter, de forma sistemática, as informações procuradas pelo pesquisador. Para LAKATOS e MARCONI (2009) apud ABREU e MORAES (2010), o questionário pode ser definido como um instrumento de coleta de dados constituído por uma série ordenada de perguntas.

O questionário é uma técnica de investigação composta por um número mais ou menos elevado de questões, sobre um determinado tema, apresentadas por escrito às pessoas. É uma ferramenta barata (pode ser distribuída para quantas pessoas o pesquisador julgar necessário), eficaz para reunir dados de uma grande população (amostra) e versátil (pode ser utilizada em qualquer

estágio da pesquisa). Além disso, o pesquisador não precisa estar presente enquanto os respondentes preenchem o questionário. Em uma entrevista, por exemplo, os respondentes (conscientemente ou inconscientemente) geram suas respostas baseados no que eles acham que o pesquisador pretende ouvir.

De acordo com OPPENHEIM (1992), os questionários não são utilizados para testar as habilidades dos respondentes. Possuem o objetivo de conhecer e/ou medir opiniões, interesses, crenças, sentimentos, expectativas, aspectos de personalidade, informações biográficas e situações vivenciadas. Portanto, deve-se pensar no questionário como um importante instrumento de pesquisa, uma ferramenta para a coleta de dados, com uma única função a desempenhar: seu trabalho é medir.

O objetivo de uma pergunta em um questionário é estabelecer uma comunicação particular. O pesquisador espera que os respondentes possuam certas informações sobre o assunto investigado. É necessário obter tais dados, com o mínimo de distorção possível. Para isso, é possível optar pela utilização de dois tipos de perguntas: as perguntas fechadas e as perguntas abertas.

Na pergunta fechada, apresenta-se um número de alternativas para uma determinada questão e solicita-se que o respondente marque a opção que ele considera mais apropriada, ou solicita-se que ele registre sua resposta em uma escala que esteja de acordo com a sua opinião. As perguntas têm que ser suficientemente bem elaboradas, para fazer com que as categorias de respostas sejam significativas. Os questionários fechados são mais fáceis e mais rápidos de responder, pois não requerem escrita. Além disso, mais questões podem ser perguntadas dentro de um intervalo de tempo. No entanto, há perda de espontaneidade e expressividade (não é possível saber o que os respondentes disseram ou pensaram através de suas próprias palavras) e possibilidade de introdução de um viés (forçar os respondentes a escolherem entre algumas alternativas e fazer com que focalizem em opções que talvez não tenham acontecido com eles).

Nos questionários abertos, solicita-se que os respondentes escrevam suas próprias respostas para as questões levantadas. Portanto, uma resposta aberta não é seguida por qualquer tipo de escolha. Essa resposta tem que ser registrada na íntegra. Este tipo de questionário pode ser útil em situações onde o pesquisador não sabe quais os assuntos são mais importantes para serem tratados. As perguntas podem ser elaboradas de maneira mais ampla, permitindo que os respondentes destaquem os assuntos que consideram mais relevantes. A maior vantagem dos questionários abertos é a liberdade que ele

fornece para os seus respondentes: é possível obter as ideias dos respondentes através da sua própria linguagem (expressada espontaneamente).

Segundo REA e PARKER (2000), a ordem na qual as perguntas são apresentadas pode afetar significativamente o estudo como um todo. Pode confundir, enviesar respostas e colocar em risco a qualidade da pesquisa. Portanto, o objetivo da pesquisa determina as perguntas que serão feitas para os respondentes. O pesquisador não tem poder sobre as pessoas e precisa convencê-las de que vale a pena participar da pesquisa. É preciso ganhar a cooperação dos participantes e motivá-los a responder as perguntas. Uma estrutura bem pensada contribui para:

- Reduzir o esforço dos respondentes;
- Assegurar que todos os temas de interesse do pesquisador sejam tratados;
- Manter o interesse dos respondentes em continuar a responder todas as questões;
- Obter o máximo de informação possível de cada respondente.

Sendo assim, o questionário deve funcionar como um diálogo, onde cada pergunta revela uma informação sobre os respondentes e provê dados para o pesquisador. A ordem das perguntas deve sugerir um passo-a-passo, indo dos itens mais fáceis para os itens mais complexos. Um primeiro princípio de estruturação é direcionar-se:

- Do mais geral para o mais específico;
- Do mais simples para o mais complexo;
- Do menos delicado para o mais delicado;
- Do menos pessoal para o mais pessoal.

Esse tipo de abordagem é conhecida como “aproximação em funil”, ou seja, o questionário começa com uma questão bem ampla, progressivamente estreita o escopo das perguntas, até chegar a alguns pontos muito específicos.

Além disso, após conseguir convencer um respondente a dar sua atenção, pelo argumento que a pesquisa trata de um assunto do seu interesse, não convém começar a interação por perguntas burocráticas (nome, sexo, idade, etc.) e também não convém começar a interação por perguntas delicadas (renda familiar, se possui alguma doença, etc.). Em outras palavras: se o respondente concorda em responder a pesquisa, porque considera a temática interessante, a primeira pergunta (e as seguintes) deve(m) tratar desta temática. Além disso, o

questionário deve apresentar perguntas agrupadas por partes que sejam interessantes para os respondentes, despertando a sua vontade de continuar.

REA e PARKER (2000) afirmam que a extensão adequada de um questionário está entre assegurar que não fique tão longo, a ponto de assustar e desmotivar os entrevistados a respondê-lo, e que cubra toda a gama de informações necessárias para aquela pesquisa. KUNIAVSKY (2003) recomenda o uso de vinte a trinta perguntas no questionário, com um tempo de vinte a trinta minutos de duração. A leitura das instruções deve durar cinco minutos. Cada pergunta deve durar de trinta (pergunta fechada) a sessenta (pergunta aberta) segundos para ser respondida. Logo, manter o questionário com vinte perguntas dá uma boa margem de segurança para o pesquisador. Questionários que tomam mais de vinte ou trinta minutos para serem respondidos, dão a sensação de serem um fardo para o respondente. Segundo JORDAN (1998), questionários devem ser curtos e concisos.

MARCONI e LAKATOS (2002) afirmam que, em média, os questionários expedidos pelo pesquisador alcançam vinte e cinco por cento de devolução. Há alguns fatores que exercem influência no retorno dos questionários: patrocinador; *layout*; extensão; carta que o acompanha solicitando colaboração; facilidades para o seu preenchimento e devolução; motivos apresentados para a resposta; classe social dos respondentes e, principalmente, o interesse pelo tema da pesquisa. Para REA e PARKER (2000), questionários longos e complexos serão um completo sucesso se os tópicos forem de interesse absoluto do respondente, como, por exemplo, se forem sobre seus filhos ou se acreditarem que suas respostas influenciarão diretamente suas vidas, com benefícios.

Junto com o questionário, recomenda-se enviar uma nota ou carta explicando a natureza da pesquisa, sua importância e a necessidade de obter respostas, tentando despertar o interesse dos respondentes, para que eles preencham e devolvam esse questionário dentro de um prazo razoável. Para estabelecer confiança, o pesquisador precisa apresentar-se e indicar com e para quem trabalha (identificando-se com uma instituição conhecida e legitimada), capturar o interesse do respondente pelo tema e expressar apreciação, ressaltando antecipadamente o quanto opiniões e experiências destes respondentes são importantes. Por isso, a etapa de introdução do questionário precisa ser persuasiva e deve conter toda a informação necessária para poder agir da maneira esperada pelo pesquisador (OPPENHEIM, 1992; REA e

PARKER, 2000). Logo, um questionário deve começar com uma introdução que apresente:

- O propósito da pesquisa;
- As instituições envolvidas;
- A qual público-alvo se dirige e porque o participante foi escolhido;
- O retorno dos resultados da pesquisa como benefícios concretos dos usuários em questão;
- A importância da participação do receptor;
- O respeito à confidencialidade da identificação do participante;
- Não há respostas corretas, apenas a opinião de cada participante;
- As instruções de preenchimento;
- O tempo de duração de preenchimento;
- As informações de contato do pesquisador.

Senão, o esforço para pedir instruções adicionais pode fazer com que a maioria dos potenciais respondentes ignore a ferramenta. É importante prestar atenção em todos esses detalhes, tanto nos questionários impressos quanto na web, pois algumas pessoas dão muita importância para estas particularidades. Caso contrário, a ferramenta pode oferecer o risco dos respondentes desconfiarem da sua credibilidade.

3.1.3. Grupo de Foco

JORDAN (1998) afirma que algumas técnicas para a avaliação da usabilidade foram desenvolvidas especialmente para a aplicação em interfaces gráficas digitais, como a co-descoberta³, o percurso cognitivo⁴ e o registro de uso⁵. Outros métodos para a avaliação da usabilidade em interfaces têm suas

³ **Co-descoberta:** envolve dois participantes que trabalham juntos para explorar um produto e/ou descobrir como determinadas tarefas são realizadas. Através da análise das verbalizações, pode-se obter um melhor entendimento das questões de usabilidade associadas com o produto (JORDAN, 1998).

⁴ **Percurso cognitivo:** o pesquisador assume o ponto de vista do usuário e trabalha em uma série de tarefas, em busca de fontes de dificuldades, erros ou inconsistências (SANTOS e MORAES, 2000).

⁵ **Registro de uso:** *softwares* que gravam as interações dos usuários, permitindo que informações sobre a interação com o produto sejam coletadas (JORDAN, 1998).

origens na psicologia, como os diários de incidentes⁶, as entrevistas, os experimentos controlados⁷ e os questionários. No caso específico da técnica conhecida como grupo de foco, JORDAN (1998) relata que a mesma é uma adaptação do marketing. Entretanto, nas pesquisas de marketing, geralmente, utiliza-se um número de oito a doze participantes no grupo de foco. Só que a investigação de problemas de usabilidade tende a envolver menos pessoas, tipicamente cinco ou seis indivíduos. Segundo SANTOS (2004), pelo fato de ser necessário um fluxo constante de discussão, o grupo deve ter uma média de seis participantes.

Segundo KUNIAVSKY (2003) apud SANTA ROSA e MORAES (2008), denominados originalmente como “entrevistas focalizadas”, a técnica de grupo de foco se desenvolveu como um método de pesquisa social nos anos 1930 e, depois, foi aprimorado como um método para melhorar a vida dos soldados durante a Segunda Guerra Mundial. Nos anos 1950, foi abraçado pelo mercado. Provavelmente, é uma das técnicas mais antigas para pesquisar a experiência do usuário.

Para ABREU e MORAES (2010), o grupo de foco é uma técnica de pesquisa qualitativa que pode reunir de dois a dez participantes qualificados, com o propósito de discutir sobre um tópico específico, por meio de questionamentos feitos e coordenados por um moderador. De acordo com JORDAN (1998), o grupo de foco é um grupo de pessoas reunidas para discutir um assunto particular. Esta discussão pode abranger, por exemplo:

- As experiências dos usuários em relação à utilização de uma interface em particular;
- Os requisitos para uma nova interface;
- As informações sobre o contexto onde realiza-se tarefas específicas;
- Os problemas de usabilidade que são associados com a utilização de uma interface.

⁶ **Diários de incidentes:** mini-questionários onde os participantes tomam notas de qualquer problema encontrado durante a utilização de um produto. Então, pergunta-se como resolveriam tal problema e como o mesmo era incômodo (JORDAN, 1998).

⁷ **Experimentos controlados:** avaliação formalmente projetada através de controles e balanços muito bem ajustados. O objetivo é remover o máximo de ruídos nos dados finais, isolando os efeitos para a execução da tarefa ser a mais clara possível. Tipicamente, a ordem das tarefas também é balanceada, minimizando possíveis efeitos de aprendizado ou transferência de conhecimento (JORDAN, 1998).

Para SANTOS (2004), através do grupo de foco procura-se obter opiniões advindas de diversos pontos de vista. É uma técnica eficaz para desenvolver conceitos e avaliar as primeiras impressões, ainda nas fases iniciais do desenvolvimento de uma interface. SANTOS (2004) também afirma que o grupo de foco pode ser utilizado para descobrir reações aos conceitos, explicitar requisitos do participante, descobrir prejuízos, além de traçar ideias sobre o que as pessoas pensam em relação a uma interface.

JORDAN (1998) afirma que o moderador segue um roteiro de assuntos que poderão conduzir o procedimento de discussão. No entanto, este roteiro não é estruturado de forma rígida, seu objetivo é permitir que os participantes prossigam em determinadas direções, conforme o desejo do grupo, assegurando que os pontos levantados serão aqueles que mais preocupam tais indivíduos. Para facilitar o grupo de foco, o trabalho do moderador é assegurar que todos os participantes tenham a mesma chance de expor suas opiniões.

Segundo JORDAN (1998), é corriqueiro que o moderador utilize um arranjo de perguntas. Este arranjo é utilizado quando um participante não é capaz de pensar em algo útil para dizer. Entretanto, é importante que o arranjo de perguntas seja simplesmente um meio de provocar mais conversações. Além disso, deve-se evitar perguntas “carregadas”, ou seja, capazes de conduzir as respostas, dando a impressão que certa interface avaliada é fácil ou difícil de se utilizar, fazendo com que os participantes sintam-se tentados a concordar com a opinião do moderador. Logo, deve-se utilizar uma linguagem neutra, onde o palavreado do moderador não dá a impressão que ele está esperando que os participantes forneçam uma resposta particular (ao contrário disso, que todas as respostas sejam genuínas). Sendo assim, as perguntas devem, simplesmente, fornecer algo concreto para os participantes discutirem e devem servir para reiniciar a conversação. Além disso, só devem ser utilizadas quando houver um problema em continuar a discussão, ao invés de serem usadas como uma maneira de redirecionar totalmente a conversa. Mesmo as perguntas neutras têm a inconveniência de poder conduzir os participantes através de assuntos que podem ser de pouca importância para eles.

Para JORDAN (1998), como em todas as técnicas que envolvem perguntas abertas, o problema em analisar as discussões a partir de um grupo de foco é o fato de interpretar o motivo de um determinado assunto não ser mencionado. Ao ser abordado, talvez um determinado problema não seja do interesse dos participantes, ou ninguém do grupo tenha pensado no assunto até aquele momento. O entusiasmo com que a conversa prossegue após a

exposição do problema pode ser um indicador deste fato. Caso o moderador continue em dúvida, o melhor a fazer é perguntar o quanto o assunto é importante.

JORDAN (1998) também ressalta que quanto mais pessoas participarem do grupo, maiores serão as chances de interação entre as pessoas. Certamente, uma das grandes vantagens do grupo de foco é o fato do comentário de uma pessoa poder gerar uma contribuição útil de outro indivíduo. Portanto, quanto mais participantes no grupo, maiores serão as chances disso acontecer. Caso haja poucos sujeitos, este efeito pode não ser alcançado. Outro fator, entretanto, é a chance de todos os participantes expressarem suas opiniões. Nesse sentido, será mais benéfico contar com poucos indivíduos, senão as pessoas terão que esperar muito tempo antes de terem a chance de falar. A consequência para este fato pode apresentar-se através de tédio, irritação ou frustração, fazendo com que alguns indivíduos sintam-se excluídos do procedimento e percam a boa vontade de contribuir. É trabalho do moderador do grupo assegurar que todos os participantes estejam envolvidos da mesma maneira, não importa o quanto sejam reservados para falar ou não.

SANTOS (2004) afirma que, de forma geral, deve-se realizar mais de um grupo de foco, pois os resultados de uma única sessão podem não ser suficientemente representativos. De acordo com EDMUNDS apud TEIXEIRA e MORAES (2004), a técnica de *focus group* pode ser compreendida como uma pesquisa qualitativa. Isto significa que seus resultados não são obtidos em porcentagem, em testes estatísticos ou em tabelas. Segundo TEIXEIRA e MORAES (2004), o grupo de foco tende a ser mais exploratório e menos estruturado do que outros métodos de avaliação de interfaces, ou seja, ao invés de prover respostas quantitativas para uma pergunta específica, os participantes de um grupo de foco fornecem respostas qualitativas, advindas da discussão em relação a um conjunto de tópicos. Isto permite que os pesquisadores capturem comentários subjetivos e avaliem considerações, percepções, sentimentos, atitudes e motivações dos participantes.

3.1.4.

Card Sorting (ou arranjo de cartões)

Segundo SANTA ROSA e MORAES (2008), a análise de *sorting* vem sendo utilizada há muito tempo nas ciências sociais, para auxiliar os pesquisadores na coleta de dados que os levem à melhor compreensão do que

as pessoas pensam sobre conceitos. Os antropólogos têm usado essas técnicas de *sorting* como meio de determinar os valores de uma cultura, as “verdades” comuns de uma comunidade e outras informações que só podem ser derivadas dos próprios membros dessas culturas / comunidades. O *card sorting* é uma técnica bem estabelecida nos campos da psicologia e da usabilidade, sendo extremamente útil para avaliar hierarquias. Na área da interação humano-computador, essa técnica é aplicada para auxiliar no projeto de interações e na arquitetura de informação, sendo utilizada para compreender como os usuários categorizam o conteúdo de uma interface.

NIELSEN e SANO (1994) afirmam que o *card sorting* é uma técnica comum de usabilidade, utilizada para descobrir o modelo mental dos usuários. Um exemplo típico da aplicação dessa técnica é a definição de estruturas de navegação, quando os usuários são solicitados, através de um grupo de cartões, a agruparem as seções da interface que fazem mais sentido para eles. Para MORVILLE e ROSENFELD (2006), o *card sorting* mostra a maneira como os usuários, geralmente, agrupam e nomeiam tarefas e conteúdos. A simplicidade do método confere muita flexibilidade para a sua aplicação. De acordo com KUNIAVSKY (2003), *card sorting* tem o objetivo de explorar o modelo mental do usuário (conhecimento, hábitos e procedimentos de organização de uma informação e categorização do seu conteúdo) no que tange a arquitetura de informação, principalmente sua estrutura de navegação. É uma técnica usada para descobrir como as pessoas organizam a informação e como elas categorizam e relacionam conceitos.

KUNIAVSKY (2003) afirma que é uma técnica rápida e fácil, logo, pode ser usada sempre que uma estrutura de informação for alterada. No entanto, o *card sorting* é mais indicado quando o pesquisador deseja investigar qual tipo de informação precisa ser organizada, antes que uma solução de organização já tenha sido implementada. Por isso, normalmente, é realizado após o produto ter seu conceito, público-alvo e funcionalidades definidos, mas antes da arquitetura de informação ou design serem desenvolvidos. É indicado em algum momento no meio do processo de pré-projeto, com o objetivo de verificar se a arquitetura da interface faz sentido para os usuários, uma vez que nem sempre o que parece ser óbvio para o projetista é para o usuário.

Segundo MORVILLE e ROSENFELD (2006) o *card sorting* tem mais utilidade quando é usado para reunir dados qualitativos. É preciso formar algumas conclusões durante os testes, enquanto os usuários falam sobre suas razões, questionamentos e frustrações. Também é possível conversar com os

participantes, para obter um entendimento melhor sobre a organização e nomenclatura do conteúdo. Um grande número de agrupamento similares e a quantidade de vezes que um mesmo cartão é colocado em uma mesma categoria, sugerem uma forte afinidade com o modelo mental dos usuários. Ao utilizar o *card sorting* através de uma abordagem quantitativa, é preciso trabalhar de acordo com os princípios básicos de metodologia científica e evitar abordagens que possam prejudicar os resultados. É importante evitar mascarar resultados qualitativos através de análises quantitativas. Em testes com apenas cinco usuários, por exemplo, os números podem não ser estatisticamente significativos. Mesmo assim, os dados serão de grande valor, pois estes cinco usuários poderão fornecer uma série de indícios sobre como a informação pode ser categorizada.

NIELSEN (2004) afirma que para realizar o *card sorting*, o indicado é aplicar a técnica com quinze usuários, para obter uma correlação aceitável. Geralmente, ao utilizar o *card sorting*, ainda não há um projeto fechado e pretende-se descobrir como as pessoas pensam a respeito de algumas questões. Há uma grande variabilidade de modelos mentais de diferentes pessoas, quanto ao vocabulário que eles usam para descrever os mesmos conceitos. É preciso coletar dados a partir de um número razoável de usuários, antes de alcançar uma estabilidade da estrutura, das preferências dos usuários e de determinar o que lhes é “comum”. Testar com cada vez mais usuários pode apresentar rendimentos decrescentes, mas mesmo assim ainda é recomendado usar três vezes mais participantes no *card sorting* do que em testes tradicionais de usabilidade. Segundo KUNIAVSKY (2003), é recomendável recrutar pessoas que se encaixam em seu público-alvo. Geralmente, um número entre quatro e dez participantes será capaz de fornecer uma boa ideia sobre como organizar a informação.

KUNIAVSKY (2003) relata que é possível utilizar nomes de seções da interface, imagens, termos que se pretende usar ou até a descrição de páginas. As palavras nos cartões devem refletir o que o pesquisador quer testar. Se o objetivo é descobrir como as pessoas organizam conceitos, é preciso escrever o conceito nos cartões com uma ou duas frases. Mas, se o objetivo é verificar como as pessoas entendem rótulos (nomenclaturas), é possível escrever as palavras sem, necessariamente, explicar o conceito delas. Recomenda-se, como um número base, uma média de cinquenta cartões por sessão. A partir de cem cartões, é melhor dividir a sessão em mais de uma aplicação da técnica.

A preparação da técnica inicia-se com uma listagem de todos os itens que se quer testar. Geralmente esses itens são retirados do diagrama de navegação do *site* ou de uma proposta de menu. Toma-se nota desses itens em pedaços de papel (ou *post-its*), transformando-os em cartões índices. Cada termo deve ficar em um cartão. No verso de cada cartão deve ser colocado, de forma discreta, um número de identificação. É comum este número ser o mesmo do termo correspondente no diagrama de navegação ou no menu proposto.

Existem duas formas de apresentar os cartões para os participantes da sessão de *card sorting*: sem a indicação de níveis, quando todos os cartões são da mesma cor (independente da hierarquia) ou com a indicação de primeiro nível, quando os cartões correspondentes aos agrupamentos principais possuem uma cor (vermelho, por exemplo) para indicar o primeiro nível da hierarquia e a partir do segundo nível (e demais níveis), outra cor (cinza, por exemplo) é utilizada. A aplicação do *card sorting* com a indicação de primeiro nível é uma das formas mais utilizadas, pois ajuda o usuário a ter um ponto de partida. Além disso, o participante também é informado que pode mudar um cartão de segundo nível para o lugar de um cartão de primeiro nível e vice-versa. Portanto, essa abordagem é uma sugestão inicial, que o usuário tem total liberdade para seguir ou não. Independente da forma de apresentação, também é importante embaralhar a pilha de cartões, antes de dar para cada participante.

Segundo MORVILLE e ROSENFELD (2006), o *card sorting* pode ser classificado em duas categorias:

- **Card sorting aberto:** os usuários têm a liberdade de alterar a nomenclatura do cartão ou das categorias. Geralmente, é utilizado para a descoberta.
- **Card sorting fechado:** utiliza-se nomenclaturas e categorias pré-definidas, sem a possibilidade do usuário modificá-las. Geralmente, é utilizado para a validação.

MORVILLE e ROSENFELD (2006) também afirmam que é preciso balancear a aplicação do *card sorting* aberto com o *card sorting* fechado, em função do que o pesquisador deseja investigar.

3.1.5. Análise da Tarefa

De acordo com MORAES e MONT`ALVÃO (2003), a base da intervenção ergonomizadora é o estudo das interações que ocorrem durante a execução da

tarefa. Utiliza-se como método a análise da tarefa. Por tarefa entende-se: o conjunto das condições de execução de um trabalho, com o propósito de alcançar um objetivo, através de um conjunto de exigências. Portanto, a análise da tarefa representa a descrição do sistema em foco, ou seja, a descrição do conjunto dos elementos que compõem esse sistema e a descrição das interações entre os elementos desse sistema, incluindo a indicação das suas disfunções. MORAES e MONT'ALVÃO (2003) também afirmam que a partir da descrição do sistema em foco, o processo de análise da tarefa visa identificar e descrever unidades de trabalho e analisar os recursos⁸ necessários para um desempenho de trabalho bem sucedido.

Para MORAES e MONT'ALVÃO (2003), a análise da tarefa é uma metodologia que faz uso de diversas técnicas, tanto para coletar informações e organizá-las quanto para fazer julgamentos, emitir diagnósticos ou tomar decisões de projeto. Para um controlador de tráfego, por exemplo, é muito importante verificar a tomada de informação em um painel sinóptico e as comunicações com a manutenção. A abordagem do ergonomista será baseada na análise da carga cognitiva e psíquica desse usuário. No entanto, para um servente de obra, os eventos a serem observados relacionam-se com as posturas e movimentação manual de materiais. Logo, a abordagem do ergonomista será baseada na análise da carga física do usuário.

MORAES e MONT'ALVÃO (2003) relatam que durante o registro comportamental das atividades da tarefa, utilizam-se observações sistemáticas, registros de comportamento (frequência, duração e sequência) e técnicas da engenharia (diagramas de fluxo).

STAMMERS e SHEPHERD (1995) afirmam que as suposições comuns sobre o que constitui uma tarefa são:

- O termo “tarefa”, geralmente, se aplica a uma unidade de atividade, dentro de situações de trabalho;
- Uma tarefa pode ser dada ou pode ser imposta a um indivíduo ou, alternativamente, pode ser levada em conta na própria iniciativa e decisão do indivíduo;
- Requer mais do que uma operação simples física ou mental para sua conclusão;

⁸ Recursos, no contexto da análise da tarefa, são aqueles trazidos pelo usuário (habilidade, conhecimento e capacidade física) e aqueles fornecidos no ambiente de trabalho (comandos, mostradores, ferramentas, procedimentos, ajudas, etc.).

- É frequentemente usada a conotação de uma atividade não trivial, ou de natureza onerosa;
- Tem um objetivo definido.

Para MORAES e MONT'ALVÃO (2003), devem ser considerados três componentes que interagem nas tarefas:

- **Requisitos da tarefa:** são os objetivos ou condições definidos pelo contexto do sistema;
- **Ambiente da tarefa:** fatores, na situação de trabalho, que limitam e determinam como um indivíduo pode desempenhar a tarefa;
- **Comportamento da tarefa:** ações que o usuário leva em consideração, inseridas nos constrangimentos do ambiente da tarefa.

Segundo JORDAN (1998), a medição da complexidade de uma tarefa é feita através do número de passos necessários para completar a mesma: quanto menos passos, mais simples é a tarefa. JORDAN (1998) também afirma que a análise da tarefa pode ser utilizada para o desenvolvimento de predições sobre:

- O quanto é fácil desempenhar uma determinada tarefa;
- O quanto é difícil desempenhar uma determinada tarefa;
- Quanto esforço é necessário para chegar ao final da tarefa.

Isto é possível porque, através deste método, a tarefa é analisada como em uma série de passos separados.

O tipo de análise da tarefa mais simples fornece os passos físicos que o usuário tem que desempenhar até o final da tarefa, como apertar botões ou acionar menus. O tipo de análise da tarefa mais complexa também demonstra os passos cognitivos envolvidos ao longo de toda a tarefa, como decidir que tipos de botão devem ser apertados ou decidir que tipos de menus devem ser acionados.

MIRANDA (2005) afirma que a análise da tarefa é considerado o método mais poderoso disponível para os que trabalham com interação humano computador, tendo aplicações em todos os estágios do desenvolvimento de sistemas. MIRANDA (2005) também relata que a análise da tarefa é uma metodologia apoiada por um número de técnicas específicas, com o objetivo de ajudar o investigador a coletar e organizar informações e utilizar as informações coletadas para emitir diagnósticos ou tomar decisões de projeto. As técnicas de análise da tarefa são tidas como:

- Uma intenção de representar a informação que será usada no projeto de um novo sistema homem/máquina;

- Uma intenção de representar a informação que será usada na avaliação de um projeto de um sistema já existente.

Isso é realizado através de análise sistemática das tarefas do usuário.

STAMMERS e SHEPHERED (1995) afirmam que o processo de análise da tarefa pode ser representado por três fases principais de atividade:

- Coleta de dados que envolvem a reunião e a documentação de várias fontes de informação sobre o sistema investigado;
- Criação de um documento de descrição da tarefa;
- Análise final da tarefa, que resulta em fontes de informação para o investigador realizar melhorias.

SANTOS (2005) afirma que o termo análise da tarefa é genérico e nomeia diversos métodos de coleta de dados, classificação de dados e interpretação de dados sobre a performance humana em situações de trabalho. SANTOS (2005) define a análise da tarefa como um processo:

- De identificar e descrever unidades de trabalho;
- De analisar os recursos necessários para um desempenho bem sucedido do trabalho.

Segundo SANTOS (2005), a análise da tarefa é um método tanto formativo, quando aplicado nas fases iniciais de desenvolvimento, quanto somativo, quando aplicado após a implementação da interface. SANTOS (2005) também afirma que a análise da tarefa é composta por dois componentes principais: a coleta de dados e a análise de dados. Ainda de acordo com SANTOS (2005), a metodologia da análise da tarefa pode ser sumarizada, de maneira geral, pelos seguintes passos:

- Obter informações sobre a função que será automatizada;
- Coletar e analisar dados obtidos com os usuários que realizam a função;
- Construir e validar um modelo junto aos usuários.

Segundo MIRANDA (2005), deve-se reunir informações sobre a tarefa através de quatro etapas:

- **Primeira etapa:** entender e identificar as necessidades para o trabalho ser automatizado;
- **Segunda etapa:** reunir membros da equipe de projeto;
- **Terceira etapa:** reunir usuários representativos;
- **Quarta etapa:** identificar e listar quais são os usuários e as tarefas.

MIRANDA (2005) também afirma que a seleção de tarefas é um trabalho difícil e extremamente importante. A frequência de uso é uma das bases para tomar decisões de projeto.

De acordo com MIRANDA (2005), devem ser selecionados uma quantidade que varia de três até cinco usuários. Também é preciso escolher usuários chave, ou seja, indivíduos que representam as “melhores práticas” - pessoas que realizam suas tarefas com eficiência e precisão. Para SANTOS (2005), é preciso escolher sujeitos representativos da população usuária. Por exemplo: Donas de casa para sistemas de compra em supermercados e bibliotecários para sistemas de informação. Os sujeitos devem ser entrevistados a fim de se conhecer seus perfis e definir o grupo de teste.

De acordo com MIRANDA (2005), a coleta de dados sobre a tarefa se faz com observações da execução da tarefa e entrevistas, com o objetivo de:

- Coletar dados sobre o ambiente de trabalho;
- Entender o modelo de trabalho do usuário;
- Aprender a linguagem do usuário;
- Coletar estatísticas sobre os casos de uso;
- Perceber os problemas, erros, gargalos e oportunidades de melhorias no processo de trabalho.

É importante realizar essa análise no ambiente de trabalho do usuário - ele tem maior facilidade para explicar suas ações.

3.1.6. Avaliação Cooperativa

A avaliação cooperativa foi desenvolvida na Universidade de York e, num primeiro momento, foi utilizada somente como uma ferramenta de pesquisa. No entanto, por se tratar de uma técnica de avaliação prática, seu potencial rapidamente se tornou aparente e a avaliação cooperativa passou a ser adotada por uma série de empresas, durante o processo de desenvolvimento dos seus produtos. Segundo MONK et al (1993) apud ABREU e MORAES (2004), a avaliação cooperativa consiste em uma técnica na qual usuários representativos, previamente selecionados, executam tarefas diretamente relacionadas à pesquisa. Ao longo da sessão de avaliação cooperativa, os usuários explicam as ocorrências para o pesquisador que, por sua vez, também faz perguntas durante a realização das tarefas. O pesquisador também permite que o usuário cometa erros e os usa para eliciar questões sobre o sistema utilizado. Comportamentos

inesperados e comentários do participante da avaliação cooperativa são vistos como sintomas de potenciais problemas de usabilidade do sistema.

De acordo com MIRANDA e MORAES (2003), o que distingue a avaliação cooperativa é a cooperação que ocorre, na medida que os participantes e o pesquisador avaliam, juntos, uma determinada interface. Neste método, os participantes são encorajados a perguntar sobre o processo de interagir com a interface, enquanto o pesquisador faz perguntas sobre o entendimento dos participantes em relação à mesma. MIRANDA e MORAES (2003) também afirmam que este método pode ser considerado como uma técnica de projeto participativo, ou seja, promove uma maneira de colocar os participantes e os pesquisadores em comunicação a respeito da própria interface. Suas principais vantagens são:

- Possibilidade de trabalhar com protótipos e simulações parciais ao invés da interface final;
- Facilidade de utilização, pois é um procedimento de baixo custo e que pode ser aplicado por qualquer indivíduo sem conhecimento prévio;
- Promove o máximo de feedback sobre como o projeto deve ser reformulado.

MIRANDA e MORAES (2003) ressaltam o passo a passo para a realização de uma avaliação cooperativa:

- **Primeiro passo:** recrutamento de participantes, que devem ser típicos da população que irá utilizar a interface;
- **Segundo passo:** preparação das tarefas, que devem ser escolhidas a fim de permitir que o participante explore as áreas da interface consideradas como relevantes;
- **Terceiro passo:** interação e registro, ou seja, tudo deve ser anotado na medida que o participante trabalha;
- **Quarto passo:** análise dos resultados.

De acordo com TEIXEIRA e MORAES (2004a), a avaliação cooperativa possibilita a união do pesquisador, da interface e do seu público-alvo. Há o envolvimento do participante com o projeto, ou seja, cada pessoa explicita suas dificuldades, suas ações e suas experiências relacionadas com a interface. TEIXEIRA e MORAES (2004a) recomendam a utilização da avaliação cooperativa em produtos que necessitam de aprimoramento técnico, em protótipos em um estágio intermediário ou em protótipos funcionando em sua plenitude. Por utilizar certas tarefas previamente formuladas, este método não é

indicado para os casos nos quais a interface se encontra em sua fase embrionária. TEIXEIRA e MORAES (2004a) também ressaltam que para um método de avaliação ser considerado como cooperativo, ele deve cumprir certas etapas:

- Recrutar 1 ou mais participantes, através de uma seleção próxima do modelo mental do usuário típico do sistema;
- Especificar certas tarefas, para os participantes utilizarem as áreas da interface consideradas como principais;
- Cada participante deve verbalizar os problemas encontrados, enquanto o pesquisador anota ou grava cada observação;
- O pesquisador, com base nas suas anotações e gravações, identifica os problemas encontrados e propõe soluções.

TEIXEIRA e MORAES (2004a) descrevem os seguintes procedimentos para a realização de uma avaliação cooperativa:

- **Etapla 1 - preparar o protótipo:** os protótipos podem ser totalmente funcionais, ou pelo menos apresentarem partes funcionais (aquelas onde as tarefas serão realizadas). No entanto, também é possível trabalhar com simulações com protótipos de papel ou *wireframes* interativos;
- **Etapla 2 - recrutar usuários:** no mínimo 1 e no máximo 5 pessoas, através de uma seleção mais próxima possível do usuário típico da interface, considerando certas características durante o processo seletivo, como o nível de experiência dos usuários, o conhecimento dos usuários sobre a tarefa a ser realizada, ou como os participantes lidam com situações que necessitam de uma conduta própria na solução de problemas diversos;
- **Etapla 3 - preparar as tarefas:** as tarefas devem contemplar áreas da interface consideradas como principais, levando em conta tarefas possíveis de serem realizadas, tarefas de real representatividade dos principais elementos e serviços contidos na interface e tarefas que sejam apresentadas em forma de lista, em um documento impresso que será entregue para o usuário no início da sessão de avaliação cooperativa;
- **Etapla 4 - cuidados especiais:** é preciso deixar o usuário à vontade, explicar o propósito da avaliação, enaltecer o foco da análise no sistema e não no usuário, mencionar o fato da sessão estar sendo registrada em áudio ou vídeo, mencionar o caráter confidencial das gravações em áudio ou vídeo, explicar que a avaliação não é um teste para o usuário (ao

contrário disso, é apenas uma forma de apresentar o sistema e as partes que serão avaliadas);

- **Etapa 5 - interagir e gravar:** cada usuário deve verbalizar os problemas encontrados (sendo constantemente estimulado para isso), enquanto o pesquisador anota suas observações (em papel e caneta) e grava as ações dos usuários (em áudio, vídeo, arquivos *log*, etc.). Todo o processo deve ser conduzido de uma maneira informal, possibilitando uma discussão franca. Além disso, deve-se permitir que os usuários cometam erros, apontem problemas e façam perguntas. O pesquisador não deve antecipar os erros para os usuários nem corrigi-los antes do tempo;
- **Etapa 6 - tabular as informações:** com base nas anotações, gravações e na própria memória do pesquisador, os problemas são identificados e soluções para eliminar tais problemas são propostas. A tabulação dos dados deve ser rápida e baseada, principalmente, nas anotações e na memória do pesquisador. Para evitar um longo tempo de tabulação, gravações em vídeo ou áudio são utilizadas somente como *backup*, ou seja, como uma prova concreta para demonstrar os problemas encontrados à terceiros;
- **Etapa 7 - apresentação dos resultados e mesa redonda:** caso seja possível organizar um segundo encontro com os usuários, recomenda-se que os resultados encontrados sejam apresentados para eles. As melhorias que serão realizadas na interface também devem ser comunicadas para os usuários. Outra alternativa (embora opcional, é altamente recomendável), apresenta-se sob a forma de uma mesa redonda, com o propósito de reunir todos os usuários para discutir, com o pesquisador, os problemas encontrados e as possíveis soluções.

3.1.7. Teste de Usabilidade

Segundo TRAVIS (2003) apud SANTA ROSA e MORAES (2008), os testes de usabilidade têm suas raízes na psicologia experimental, na qual se pedem que os participantes executem tarefas bem determinadas e se faz uma análise dos resultados. SANTA ROSA e MORAES (2008) também afirmam que o teste de usabilidade é um método, empregado na ergonomia e na interação humano-computador, para testar e avaliar a usabilidade de produtos e sistemas, a partir da observação dos usuários durante a interação. O foco do teste de usabilidade

está no comportamento observável - o que os usuários fazem mais até do que dizem que fazem.

De acordo com RUBIN (2008) apud ABREU e MORAES (2010), o termo “teste de usabilidade” é, frequentemente, utilizado de forma indiscriminada para se referir a qualquer técnica utilizada para avaliar um sistema. No entanto, o teste de usabilidade é um processo no qual participam pessoas representativas do público-alvo para avaliar o grau em que uma interface atende a critérios específicos de usabilidade.

DUMAS e REDISH (1999) afirmam que os projetistas e desenvolvedores devem ser lembrados que o objetivo principal de um teste de qualidade de *software* é revelar os *bugs* (erros) antes do lançamento do produto. Os objetivos do teste de usabilidade são similares aos do teste de qualidade de *software*, pois se trata de um procedimento que visa desvendar os problemas de usabilidade mais sérios de um produto. Sendo assim, a meta é criar tarefas para sondar áreas que tenham potenciais problemas de usabilidade.

- Tarefas que sondam problemas potenciais de usabilidade;
- Tarefas a partir da experiência e dos interesses do pesquisador;
- Tarefas que os usuários devem executar no produto.

O foco do teste de usabilidade é encontrar problemas reais no produto, ou no processo utilizado para desenvolver o produto. Problemas reais são aqueles que causam dificuldade para usuários, quando estes interagem com o produto em sua casa, trabalho, etc. Muitos produtos são complexos demais para que todos os seus potenciais usuários sejam testados. Mesmo no caso de produtos mais simples, acontecem tantas coisas durante um teste de usabilidade que, se os objetivos e os pontos de interesse não estiverem bem definidos, aspectos importantes podem ser perdidos. A definição dos objetivos e dos pontos de interesse são critérios fundamentais e facilitam a seleção dos usuários do teste, das tarefas e dos demais aspectos do planejamento.

DUMAS e REDISH (1999) também relatam que em testes de usabilidade, geralmente, é possível reunir uma grande quantidade de dados sobre um pequeno número de participantes. Por isso, o teste de usabilidade é um processo de amostragem. É complicado testar todas as tarefas realizadas pelo usuário em um produto. Uma boa tarefa é toda aquela que tem o potencial para descobrir um problema de usabilidade. Os profissionais que desenvolvem o produto sempre têm noção das áreas onde podem existir potenciais problemas de usabilidade.

Segundo DUMAS e REDISH (1999), o teste deve ser realizado por um especialista, mas essa pessoa tem que trabalhar com os demais membros da equipe. O ideal é contar com três pessoas no time, ou seja, dois especialistas em usabilidade e uma pessoa técnica. O teste pode ser conduzido por mais ou menos pessoas, entretanto, um número muito grande de pessoas pode atrapalhar. Cada profissional tem uma contribuição a dar. Por exemplo:

- O especialista sabe conduzir o teste e conhece os possíveis problemas;
- O desenvolvedor conhece o produto e pode ajudar a definir as tarefas;
- O *help desk* conhece os principais problemas de produtos similares;
- O pessoal de marketing sabe quais são os usuários que devem ser chamados para o teste.

O planejamento de um teste de usabilidade envolve uma série de passos, como, por exemplo:

- Definir objetivos e pressupostos do teste;
- Definir quem serão os participantes;
- Recrutar os participantes;
- Selecionar e organizar as tarefas do teste;
- Criar o cenário⁹ das tarefas;
- Decidir como identificar, medir e classificar os problemas de usabilidade;
- Preparar os outros materiais do teste (roteiros, *check-lists*, planilhas, etc.);
- Preparar o ambiente do teste;
- Preparar a equipe que conduzirá o teste;
- Realizar um teste piloto e fazer os ajustes necessários.

Segundo DUMAS e REDISH (1999), usabilidade significa fazer usuários produtivos e felizes. Sendo assim, ao estabelecer critérios quantitativos para o teste de usabilidade, é preciso manter o foco no usuário e não no sistema ou produto. Técnicas como análise da tarefa, entrevistas, *focus group*, testes de usabilidade em momentos anteriores, além de outros tipos interações com o usuário, podem ajudar o pesquisador a estabelecer os critérios a serem medidos

⁹ Os cenários são usados para informar aos participantes o que o pesquisador espera que eles façam durante o teste. Esta técnica descreve as tarefas de uma maneira que diminui a artificialidade do teste. Por exemplo: “você acabou de retornar ao seu trabalho após as suas férias. Verifique quantos e-mails você tem para ler e caso tenha recebido alguma mensagem do Sr. Green, o Vice-Presidente da empresa, leia antes das demais”.

no teste de usabilidade. Medidas de desempenho, imaginar o ponto de vista do usuário e o senso comum permitem a seleção desses critérios.

3.1.8. *Eye Tracking*

Eye tracking é uma técnica que, com a ajuda de um equipamento, consegue seguir o olhar do usuário e verificar exatamente para quais pontos ele está olhando, além de gravar a ordem em que ele visualizou cada área e o tempo em que ele fixou o olhar em cada ponto. Tudo isso de uma forma não intrusiva. Desta forma, com a utilização de mapas de calor, vídeos, gráficos e métricas relacionadas ao tempo de visualização, é possível analisar os dados sobre a ordem de visualização das pessoas e de áreas mais vistas.

JOSEPHSON (2004) relata que há cerca de cem anos a técnica de *eye tracking* tem sido utilizada por psicólogos, como uma forma de medir a relação da cognição com a atenção. JOSEPHSON (2004) também afirma que desde 1950 e, especialmente, desde a década de 1970, a pesquisa com uso de *eye tracking* tem sido usada para estudar como os consumidores olham para anúncios impressos, catálogos, etiquetas de advertência de tabaco e álcool, rótulos nutricionais dos alimentos, páginas amarelas, anúncios de televisão e, mais recentemente, *banners* em páginas da *web*. A técnica também tem sido utilizada para examinar a atenção visual dada para produtos expostos nos supermercados e no processo de escolha do consumidor em geral. RUSSO (1978) foi um dos primeiros pesquisadores a recomendar o uso de *eye tracking* para estudar a publicidade, indicando a utilização em conjunto com protocolos verbais em ambientes de consumo.

FITTS, JONES e MILTON (1950) apud RUSSEL (2005) afirmam que a técnica de *eye tracking* tem sido utilizado em testes de usabilidade por muitos anos, começando principalmente com avaliações do *design* de cabines de pilotagem de aviões. Esses primeiros estudos foram importantes no estabelecimento de algumas hipóteses sobre as relações entre os dados do movimento dos olhos e da atividade cognitiva dos usuários. A frequência de fixações foi relacionada com a importância do controle, enquanto que a duração das fixações foi relacionada com a facilidade de interpretação da informação.

Segundo GOULD e ZOLNA (2010), desde a sua criação, o *eye tracking* tem sido empregado:

- Por cientistas cognitivos para estudar a leitura, aprendizagem e atenção;

- Por profissionais de marketing para examinar a eficácia de projetos de publicidade e embalagem;
- Por engenheiros e ergonomistas para a construção do interior de automóveis e cabines de avião.

Estas e outras disciplinas têm tido grande sucesso de forma a considerar o *eye tracking* como uma técnica de pesquisa comportamental para obter informações também sobre as interações do usuário com as máquinas.

Segundo JOSEPHSON (2004), no início, os equipamentos eram muito rudimentares. Mas de acordo com NIELSEN e PERNICE (2010), os equipamentos mais novos (dos anos 2000 para cá) utilizados para realizar *eye tracking* lançam um feixe de luz infravermelha sobre os olhos dos usuários. Esse comprimento de onda é melhor refletido pela retina do que pelo resto do olho, uma vez que a retina absorve a luz visível e reflete a luz infravermelha. Isso ajuda o equipamento de *eye tracking* a identificar a posição das pupilas, ou seja, a identificar para onde os usuários estão olhando.

Hoje, existem vários equipamentos diferentes à venda no mercado. Mas todos apresentam um alto custo de aquisição. Encontra-se um equipamento deste tipo na faixa entre vinte e trinta mil dólares. Basicamente, há três modelos: em formato de monitor, em forma de óculos ou uma espécie de caixa portátil, que pode ser utilizada em diferentes cenários.

Como ainda não há fabricantes no Brasil, é preciso fazer uma importação para realizar estudos com esse tipo de tecnologia, o que aumenta ainda mais os custos. Por outro lado, já existem universidades que utilizam estes equipamentos para pesquisa científica na área de oftalmologia, e algumas empresas de pesquisa já fazem aluguel do equipamento ou realizam estudos com a utilização desta tecnologia. Porém, existem poucas empresas privadas no Brasil com esse tipo de equipamento, para a realização de pesquisas *in-loco*. No entanto, a empresa escolhida para a realização do estudo de caso desta pesquisa de mestrado, a Globo.com, adquiriu em 2010 um equipamento próprio para realizar *eye tracking*. A proponente desta pesquisa desenvolveu o projeto do novo laboratório de usabilidade da empresa e é a responsável por todas as pesquisas de usabilidade da Globo.com, inclusive, com a utilização do equipamento X60 da Tobii, para realização de *eye tracking*. Por conta disso, tornou-se possível a utilização desta tecnologia nesta pesquisa de mestrado.

ROSS (2009) explica que o equipamento de *eye tracking* detecta quando a fóvea de uma pessoa fixa em um ponto, além de detectar as movimentações entre um ponto e outro. A fóvea é uma pequena região da retina que é

responsável pela visão detalhada, com alta resolução. Fora da fóvea, a acuidade visual diminui muito. Os olhos se movem constantemente, de forma bem rápida. Esse movimento é chamado de “sacada”. Entre uma sacada e outra, há breves paradas em alguns pontos, que chamamos de “fixações”. Nessas paradas, a informação visual é captada através da fóvea. A área fora da fóvea (parafoveal) e a visão periférica determinam onde fixar o próximo ponto.

É possível observar mais claramente esses pontos de fixação e sacadas no mapa gerado com os dados do *eye tracking*, chamado de “percurso do olhar” (*gaze plot*, em inglês), que será detalhado adiante. A figura 3.1. mostra um exemplo do percurso do olhar de um usuário ao visualizar uma página do Receitas.com, objeto de estudo desta pesquisa de mestrado. Na imagem, os círculos representam os pontos de fixação do olhar, onde a fóvea do participante parou, brevemente, para absorver a informação. Isso aconteceu quando a “imagem” passou pela pupila e se formou na parte de trás dos olhos. As linhas que ligam os círculos são as sacadas, ou seja, os movimentos entre as fixações.

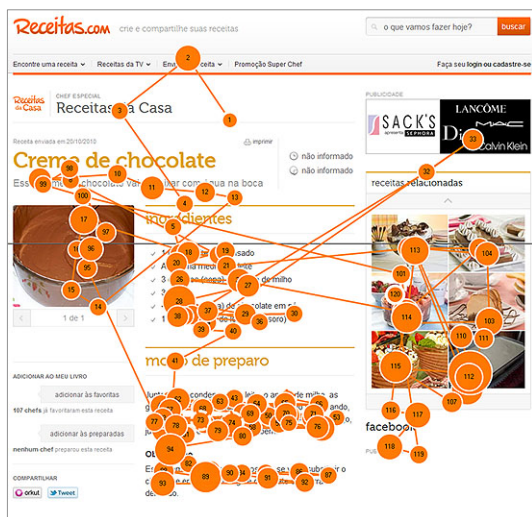


Figura 3.1. - Exemplo de fixação e sacada, onde cada círculo representa uma fixação e cada linha uma sacada.

Segundo JOSEPHSON (2004), a pesquisa com o uso de *eye tracking*, sobre os movimentos oculares na percepção de uma cena e na leitura, fornece dados para ajudar a medir a quantidade de atenção visual, bem como a relação com o nível de processamento cognitivo que ocorre. HOFFMAN (1998) e RAYNER (1998) apud JOSEPHSON (2004) relatam as informações que foram acumuladas através de uma vasta quantidade de pesquisas. Sabe-se que os olhos se movem continuamente, de duas a cinco vezes, a cada segundo que absorvemos a informação visual. Os movimentos oculares consistem em

fixações durante as quais o olho permanece relativamente imóvel por cerca de cem até trezentos milissegundos, separados por movimentos rápidos, como sacadas, cujo valor médio é de três a cinco graus de ângulo visual e com a duração de quarenta a cinquenta milissegundos. Descobriu-se também que a localização da fixação do olho é um bom indicador da atenção visual para estímulos complexos, porque:

- Pouca ou nenhuma informação complexa pode ser extraída durante as sacadas;
- A atenção na área da fóvea é mais eficiente do que a atenção fora da fóvea (parafoveal);
- A acuidade visual deteriora-se rapidamente fora da fóvea.

Apesar das vantagens descritas por JOSEPHSON (2004), ROSS (2009) trata como restrição o fato do equipamento só capturar a visão “nítida”, que é a que se forma na fóvea através das fixações. Como a visão periférica não é capturada, e esta representa 98% do campo visual, trata-se de uma limitação relevante, já que é a visão periférica que ajuda a decidir o que vai ser fixado na sequência (as pessoas podem ver as coisas em uma tela, sem fixar diretamente sobre elas). Por exemplo: é possível visualizar os *banners* no lado direito da página, utilizando a visão periférica. As pessoas sabem o suficiente sobre a sua posição no lado direito da página, o seu aspecto visual, além de usarem suas experiências anteriores na *web* para identificá-los e, muitas vezes, optarem por não fixar o olhar sobre eles. Como os dados de um *eye tracking* não mostram a visão periférica, é preciso tomar cuidado ao interpretar as imagens geradas, pois elas não mostram tudo que o usuário viu. Na verdade, mostram somente as fixações, ou seja, o que ele viu com mais detalhe, em alta-resolução.

Muitos questionam se realmente é possível provar a relação entre a atenção e as fixações. ROSS(2009) ressalta que fixações não representam, necessariamente, atenção ou compreensão. O *eye tracking* mostra apenas os pontos fixados. Ele não revela nada sobre os processos de alto nível de atenção e compreensão. Só porque os olhos de uma pessoa se fixam em alguma coisa, isso não significa que essa pessoa está conscientemente atenta à esse elemento. ROSS (2009) também acrescenta que a fixação não comunica significado. O *eye tracking* mostra apenas para onde o usuário olhou com mais detalhe e o número de vezes que olhou para cada ponto. Isso quer dizer que o equipamento indica “onde” mas não o “por quê”. Por exemplo: se um elemento recebeu muitas fixações, ou fixações mais longas, isso pode indicar tanto atração por esse elemento quanto dificuldade para entendê-lo. Portanto, é

importante que os dados não sejam utilizados sozinhos, uma vez que esses dados do *eye tracking*, por si só, têm pouco significado e a técnica sozinha não é suficiente. A observação cuidadosa da tarefa e a discussão com os participantes são necessárias para entender seus modelos mentais. A aplicação da técnica de *eye tracking* fornece dados complementares para ajudar neste entendimento. Sendo assim, a técnica de *eye tracking* exige interpretação, a observação cuidadosa e uma discussão com cada participante, para permitir um melhor entendimento do contexto, proporcionando a interpretação correta dos dados extraídos. RUSSO (1978) apud JOSEPHSON (2004) examinou cinco metodologias aplicáveis à pesquisa do consumidor, dentro e fora do laboratório, e chegaram a conclusão que nenhum método é tão diferente de uma análise das fixações quanto os protocolos verbais. Por isso, indica-se uma conversa posterior à aplicação da técnica de *eye tracking*, como forma de complemento.

Para GOULD e ZOLNA (2010), a atenção é um construto psicológico definido como um elemento do funcionamento cognitivo, em que o foco mental é mantido sobre um assunto específico, objeto ou atividade. O movimento dos olhos está intimamente ligado à atenção: “você não pode mover seus olhos sem mover a atenção, mas você pode mover a atenção sem mover os olhos”. Mesmo quando os olhos e a atenção não estão em sintonia, logo que se move a atenção para uma nova posição, os olhos vão querer seguir também. Portanto, num contexto de tentar compreender o que chama a atenção das pessoas em um *site*, a atenção e o foco dos olhos são quase indistinguíveis. Dessa forma, a interpretação dos dados do *eye tracking* deve ser realizada com a compreensão de que os olhos não refletem diretamente a atenção, mas eles são a nossa melhor janela para o que os usuários estão prestando atenção.

GOULD e ZOLNA (2010) citam, como exemplo da questão da atenção *versus* o olhar, um cenário sugerido por Jared SPOOL, sobre o problema de encontrar uma garrafa de *ketchup* na geladeira. A pessoa pode estar olhando diretamente para a garrafa, no entanto, por muitas razões, pode não estar prestando atenção suficiente para perceber que essa garrafa está ali. O *eye tracking* indicaria que essa pessoa viu o *ketchup*, mesmo que ela relate que não tenha visto. O risco deste cenário é concluir, erroneamente, que o usuário foi bem-sucedido em sua tarefa, quando na verdade não foi: o usuário fixou seu olhar sobre a garrafa de *ketchup*, mas não tirou essa garrafa da geladeira e ainda relatou não tê-la visto. Isso indica que ele realmente não a enxergou, mesmo com a localização dessa garrafa na sua área de visão. Mas, quando combinamos com outros dados, esse risco aparente pode ser convertido em

uma oportunidade para derivar uma observação importante, sugerindo aos pesquisadores outros fatores que podem contribuir para a garrafa não despertar a atenção ou compreensão do usuário como, por exemplo, uma garrafa opaca, um rótulo pouco claro, etc.

DUCHOWSKI, (2003) e HYÖNÄ, RADACH e DEUBEL (2003) apud RUSSEL (2005) ressaltam que, nos últimos anos, o *eye tracking* foi adaptado para estudos de usabilidade envolvendo *websites*. Os profissionais de usabilidade e grupos de pesquisa independentes vêm tentando identificar as contribuições específicas da técnica para avaliar a usabilidade de um *site*. Muitos desses estudos envolvem a identificação de que tipo de elementos específicos fazem com que atenção dos usuários seja direcionada para diferentes partes da interface, ou áreas de interesse (*areas of interest*, em inglês, cuja sigla é AOI). Os dados relativos às AOIs podem ser comparados de diferentes maneiras, incluindo o número, a ordem e a duração de fixações para cada AOI.

No entanto, algumas pessoas da área ainda questionam se a utilização de *eye tracking* em estudos de usabilidade realmente traz algum benefício que justifique o alto custo do equipamento. Como resposta à essas dúvidas, ROSS (2009) relata o estudo “Precisamos de *eye trackers* para dizer onde as pessoas olham?” (*Do we need eye trackers to tell where people look?*), apresentado no CHI 2006 por JOHANSEN e HANSEN, cujo objetivo era descobrir se especialistas em usabilidade seriam capazes de prever como seria o percurso do olhar dos usuários em uma determinada interface, utilizando o teste de usabilidade em conjunto com o *eye tracking* para verificar se os resultados seriam diferentes dos que podem ser obtidos através de um teste de usabilidade convencional. Foi solicitado para os participantes relatarem tudo o que viram na interface, utilizando apenas sua memória. Descobriu-se que os participantes se lembravam de 70% dos elementos que eles realmente tinham visto, mas não foram muito bons ao relatarem a ordem em que tinham olhado para cada um desses elementos. Os pesquisadores também pediram para os especialistas em usabilidade preverem os movimentos dos olhos dos participantes na mesma interface. A descoberta foi de que os especialistas previram apenas 46% dos elementos que os usuários viram. Além disso, apenas 36% foram capazes de prever as três primeiras áreas da tela para onde os participantes olharam. O estudo concluiu que os especialistas não podem prever, com precisão, para onde as pessoas vão olhar, assim como os participantes não podem informar, com confiança, para onde eles olharam e, principalmente, em qual ordem. Desta

forma, ROSS (2009) complementa que o *eye tracking* fornece dados adicionais valiosos sobre as áreas para onde os participantes olharam durante as tarefas.

NIELSEN e PERNICE (2009) afirmam que em testes de usabilidade tradicionais, sem *eye tracking*, os usuários fornecem muitas pistas que ajudam a avaliar o *design*, como o tempo que uma pessoa passa em uma página, a ação que ela executa por lá, o que lê em voz alta, onde passa o mouse, se sorri ou se faz caretas, o que comenta e etc. Portanto, tais comportamentos compreendem os aspectos mais importantes sobre o aprendizado a respeito de usabilidade. Mas os autores também relatam que o uso do *eye tracking* fornece uma outra dimensão, adicionando um nível a mais de informação para estudos de usabilidade ao verificar o que as pessoas estão olhando. Seguir os olhos dos usuários ao percorrerem a tela é como estar na cabeça deles, pensando com eles. Se não servir para mais nada, pelo menos serve para prender a atenção do observador com mais facilidade.

Para NIELSEN e PERNICE (2009), a tecnologia de *eye tracking* permite compreender além dos cliques e pausas. É possível aprender o que chama a atenção e por quê. Ao estudar o que as pessoas olham e não olham, o pesquisador é capaz de reunir critérios mais detalhados sobre como funciona o comportamento do usuário em interfaces usáveis e não tão usáveis assim. Comportamentos como revisão exaustiva (onde as pessoas olham repetidamente em áreas que deveriam ser úteis, mas não são) e cortes do olhar em áreas que passam a ser desconsideradas (onde as pessoas, propositadamente, passam a desconsiderar áreas do *site* em determinados momentos) ficam muito evidentes quando há um procedimento para acompanhar o olhar do usuário durante o uso de um produto. Pode-se observar quais os itens da interface que desviam a atenção do usuário sem, durante o projeto, haver a intenção de fazê-lo. Também é possível observar quais são as áreas mais “quentes” da tela, ou seja, que realmente chamam a atenção do usuário.

No entanto, para NIELSEN e PERNICE (2009), realizar pesquisas de usabilidade já exige um bom planejamento e experiência. Mas para realizar estudos de *eye tracking* em usabilidade, acrescenta-se outro plano na preparação, facilitação e análise. A tecnologia de *eye tracking* engloba um teste de usabilidade padrão e o torna mais difícil, mais demorado e mais caro. Conduzir ou analisar resultados de um estudo de forma equivocada é pior do que não fazer nenhuma pesquisa. Infelizmente, a tecnologia de *eye tracking* torna muito fácil obter resultados enganosos, que pode fazer acreditar em algo sobre o projeto que não é verdade. Os autores ressaltam que, enquanto vemos

muitos benefícios para a pesquisa com *eye tracking*, bons estudos de usabilidade não necessitam de soluções de alta tecnologia. Alguns dos melhores e iterativos testes são feitos somente com papel e lápis, porque quanto mais rápido for a execução de um estudo, mais cedo será possível ajustar a direção de projeto.

ROSS (2009) ressalta que o *eye tracking* não é apenas um *software* e um *hardware* que o pesquisador utiliza para gravar sessões de testes de usabilidade. Para usá-lo de forma eficaz, antes, é preciso fazer alterações no processo padrão de testes de usabilidade. Para obter dados precisos, é importante que o usuário realize a tarefa sem interrupções, para que o seu olhar não seja desviado ao ter que responder a uma pergunta do moderador, “contaminando” os dados de sua visualização. Inclusive, há uma certa polêmica entre a recomendação do uso ou não da técnica “pensar alto” (*think aloud*, em inglês) junto com testes de usabilidade que utilizam o *eye tracking*. Na técnica de *think aloud*, o usuário é incentivado a falar tudo que ele estava pensando ao interagir com a interface. Há quem defenda que isso deve ser mantido, inclusive com *eye tracking*. Mas, a maioria acredita que a junção das três técnicas pode prejudicar os dados do *eye tracking*. ROSS (2009) e NIELSEN e PERNICE (2010) recomendam que o usuário não seja interrompido durante a tarefa e que deve-se usar o *think aloud* de forma retrospectiva, após a tarefa, quando o usuário é solicitado a descrever o que ele fez e pensou durante a tarefa. Neste momento, o pesquisador (ou moderador) também pode fazer perguntas. Ao recorrer à técnica “pensar alto” depois da tarefa (ao invés de recorrer à mesma durante a tarefa), utiliza-se o nome “falar alto” (*talk aloud*, em inglês), uma espécie de descrição (também conhecida, em inglês, como *debriefing*).

Para ROSS (2009), descrever a tarefa após ela ter sido completada pode não trazer o mesmo grau de detalhes do que falar enquanto ela acontece, uma vez que a descrição posterior depende da memória, que pode ser falha. GOULD e ZOLNA (2010) acrescentam que um dos mais valiosos usos dos dados gerados com o *eye tracking* é a informação que complementa outras observações. Na pesquisa em ciências sociais, dá-se o nome de “triangulação” à esse cruzamento de informações resultante da aplicação de diferentes técnicas. Isso é importante para prevenir erros e interpretações equivocadas, que podem surgir através das “fraquezas” de um único método. A combinação dos testes de usabilidade e do “pensar alto” com o *eye tracking* pode ajudar a completar a pesquisa, cruzando o que os usuários dizem com o que vêem. Assim, os dados mais valiosos, que podem ser colhidos através do *eye tracking*, são aqueles

relacionados com o comportamento, que não podem ser observados a olho nu ou através de uma conversa detalhada, mas que podem ser complementados por esses tipos de observações cruzadas.

STEVENS et al (2006) apresentaram a técnica chamada “protocolo de experiência posterior com *eye tracking*” (*post experience eye tracking protocol*, em inglês, cuja sigla é PEEP). Eles realizaram um estudo que comparava o uso da técnica de *think aloud*, durante a realização de uma tarefa, com a utilização de *eye tracking* com a descrição posterior, utilizando o PEEP. A técnica consiste em, após a tarefa, exibir para o usuário o vídeo gerado pelo *eye tracking* com os dados de suas visualizações, para que ele possa fazer uma descrição mais precisa, de forma a não depender somente da sua memória. Para os autores, a chance do usuário esquecer algo nessa descrição posterior é minimizada pelo fato do vídeo ser exibido para ele. Além disso, segundo o estudo, o *eye tracking* com a utilização simultânea do *think aloud* sofre grande influência, se comparado com os dados sem a utilização do *think aloud*. Desta forma, a partir da apresentação deste estudo, muitos pesquisadores começaram a utilizar a técnica PEEP após a realização de *eye tracking*.

NIELSEN e PERNICE (2009) classificam essa conversa, após o término das tarefas e gravação do *eye tracking*, como “retrospectiva”. Os autores comentam que existem algumas formas de fazer isso, inclusive mostrando o vídeo da gravação que acabou de ser feita para o usuário (como no método PEEP, descrito anteriormente). Porém, NIELSEN e PERNICE (2009) ressaltam que, na maior parte dos casos, preferem não aplicar tal técnica. Os autores acreditam que a simples conversa, ou preenchimento de um questionário posterior, é capaz de proporcionar as informações necessárias para os esclarecimentos, sem correr o risco do usuário, ao assistir o vídeo, “fabricar” pensamentos, tentando explicar algo que, durante a navegação, não tinha parado para pensar a respeito. Na maior parte do tempo em que as pessoas navegam na internet, elas não ficam pensando muito sobre o que e como fazer tal coisa. Portanto, a utilização de tal abordagem pode ser muito demorada e perigosa. Caso seja feita, deve ser realizada somente no final da sessão, após o pesquisador ter feito as perguntas que precisa e conseguir explicações, antes mesmo de apertar o botão “play” do vídeo.

É importante ter em mente que o fato de exibir um vídeo com os dados coletados durante a navegação do usuário, para que ele possa descrever o que estava pensando naquele momento, pode influenciá-lo, uma vez que esse usuário pode não ter registrado algo em sua memória e se sentir na obrigação

de explicar algo que ele não se lembra, mas que o vídeo mostra que ele “olhou”. Em função disso, a proponente desta pesquisa de mestrado tem algumas restrições quanto à mostrar o vídeo gerado pelo *eye tracking* com os dados das visualizações dos usuários, solicitando que eles façam comentários. No entanto, tudo depende do que está sendo estudado e em quais condições. Para controlar melhor as variáveis e não correr o risco de influenciar o usuário ao exibir o vídeo, neste estudo realiza-se uma entrevista posterior ao uso do *eye tracking*, sem exibir o vídeo para os participantes.

NIELSEN e PERNICE (2010) também alertam que é preciso ter muito cuidado ao desenvolver as tarefas para um estudo com *eye tracking*. Em muitas pesquisas, os usuários devem olhar para uma ou duas páginas, o que não reflete uma navegação real. Esse tipo de abordagem pode fornecer dados que não são compatíveis com a realidade. Portanto, o indicado é apresentar tarefas que simulem a navegação entre várias páginas, não limitando a experiência de uso real, mesmo que o objetivo seja estudar somente uma parte da interface. Algumas vezes, ao realizar *eye tracking*, tarefas mais abrangentes podem simular melhor a realidade de uma navegação na *web* do que tarefas muito específicas. Sendo assim, ao analisar os mapas de calor que são gerados, é fundamental levar em consideração a tarefa dada ao usuário. Pode ser que o mapa de calor de uma página com uma determinada tarefa seja muito diferente do mapa desta mesma página, caso o usuário esteja realizando outra tarefa.

3.1.8.1.

Tipos de análise

Os dados gerados com o uso de *eye tracking* permitem uma série de análises, algumas mais aprofundadas e complexas e outras não tão detalhadas, que fornecem apenas uma visão mais abrangente. Apresenta-se a seguir os tipos de análise mais utilizados na área de avaliação de usabilidade. Os outros tipos, que dependem de mais precisão científica, muitas vezes utilizados até na área de oftalmologia, serão apenas listados, como uma forma de indicar a sua existência (em caso de necessidade de mais informações sobre estes outros tipos de análise, recomenda-se o material da pesquisadora Sheree JOSEPHSON - 2004).

Na área de usabilidade, as análises mais comuns a partir dos dados gerados com o uso do *eye tracking* são:

- Mapa de calor (*heat map*);

- Mapa de calor invertido (*gaze opacity heat map*);
- Percurso do olhar (*gaze plot*);
- Mapa com fixações próximas, agrupadas em áreas (*cluster*);
- Análise por áreas de interesse (*area of interest* ou AOI);
- Análise do vídeo de cada usuário;
- Análise através da reunião de fixações em um único mapa, para verificar elementos fixos.

Além dos tipos de análise listados acima, é possível realizar outras análises mais aprofundadas a partir dos dados gerados com o uso do *eye tracking*, como:

- Número de fixações;
- Localização das fixações;
- Duração das fixações;
- Tempo acumulado das fixações;
- Fixações no alvo;
- Número de sacadas;
- Amplitude das sacadas;
- Matriz de transição;
- Todas as fixações antes e depois do alvo;
- Fixações por áreas.

NIELSEN e PERNICE (2010) relatam que existem três formas principais de visualizar os resultados de *eye tracking*. A melhor abordagem é assistir aos vídeos do percurso do olhar do usuário em câmera lenta. Mas isso pode ser muito demorado, além de não permitir uma boa visualização dos resultados. Para isso, os mapas estáticos que são gerados podem ser mais eficientes, como o mapa de calor (*heat map*, em inglês) e o mapa do percurso do olhar (*gaze plot*, em inglês). Porém, ao observar esses mapas, é preciso ter em mente que eles representam o movimento dos olhos dos usuários ao longo do tempo. Portanto, o comportamento real ao olhar para uma página, cujos olhos se movimentam com muita rapidez, é sensível ao tempo. Sendo assim, ao analisar a usabilidade, também deve-se considerar a dimensão de tempo como um dos componentes mais importantes de um estudo.

3.1.8.2. Mapa de calor (*heat map*)

O mapa de calor é um mapa gerado com os dados das visualizações na página. A escala de cor pode ser configurada para as cores que o pesquisador desejar. Pode-se gerar o mapa de duas formas:

- Medindo a quantidade de pontos visualizados em cada área;
- Medindo o tempo total de cada ponto.

A diferença é que, ao gerar o mapa pela quantidade de pontos, o sistema trata da mesma forma os pontos com mais e menos duração. No entanto, ao gerar o mapa pelo tempo, ele dá mais peso às fixações com mais tempo e menos peso às fixações com menos tempo. Na figura 3.2., o mapa de calor da página exibe as áreas visualizadas por um usuário, onde as áreas em vermelho foram vistas por mais tempo e as áreas em amarelo e em verde, respectivamente, foram vistas por menos tempo.

Além de gerar o mapa de calor com os dados de um usuário, também é possível gerar um mapa de calor com os dados de todos os usuários que visualizaram essa página.

Existe uma limitação técnica nesse tipo de gráfico, e dos demais que serão apresentados a seguir. Se a página apresentar elementos dinâmicos (como menus *drop-down*, galerias de foto, etc.), que são abertos sobre a página sem gerar um outro endereço *web* (URL), a ferramenta não é capaz de diferenciar os diferentes estados da página e assume que todas as visualizações foram para uma página estática. Sendo assim, deve-se ter muito cuidado ao observar esses mapas, caso componentes dinâmicos estejam presentes na página. Nessas situações, é melhor não gerar os mapas de calor e analisar os resultados somente pelos vídeos.

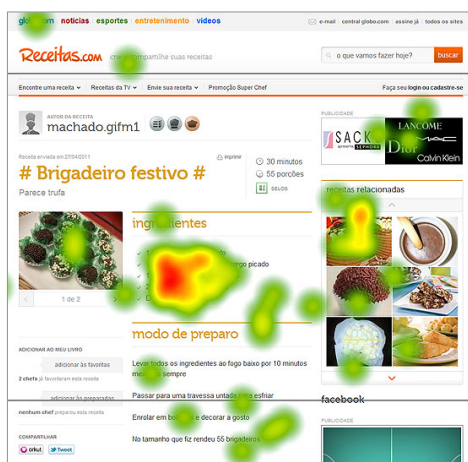


Figura 3.2. - Mapa de calor com as visualizações de uma pessoa em uma página.

3.1.8.3.

Mapa de calor invertido (*gaze opacity heat map*)

O mapa de calor invertido tem as mesmas características do mapa de calor, só que ele inverte o fundo, deixando-o opaco, e apresenta “buracos” nas áreas vistas pelo usuário. Da mesma forma que o mapa de calor, também pode ser gerado a partir da reunião das visualizações de todos que navegaram na página, além de apresentar as mesmas limitações técnicas para a geração de dados relacionados aos elementos dinâmicos da página. (figura 3.3.).

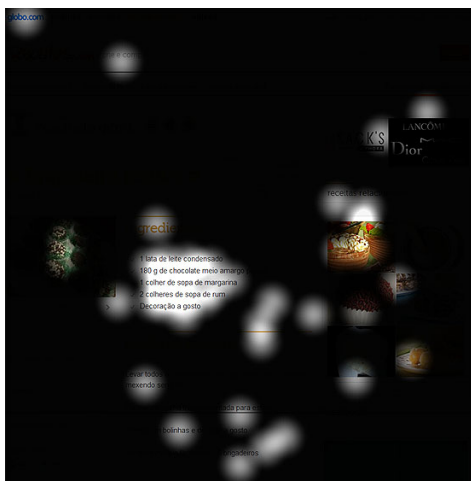


Figura 3.3. - Mapa de calor invertido com as visualizações de uma pessoa em uma página.

3.1.8.4.

Percurso do olhar (*gaze plot*)

A figura 3.4. apresenta o mapa com o percurso do olhar de um usuário ao longo de uma página. Na imagem, os círculos representam os pontos de fixação do olhar, os números indicam a ordem das fixações e o tamanho de cada círculo indica a duração de uma fixação (círculos maiores indicam fixações mais longas: pontos vistos por mais tempo). As linhas que ligam os círculos são as sacadas, ou seja, os movimentos dos olhos entre as fixações. Cada usuário recebe uma cor no mapa, para representar seus pontos visualizados. A figura 3.4. exibe apenas o percurso do olhar de um usuário, enquanto a figura 3.5. reúne os dados de todos os usuários que visualizaram a mesma página, sendo que cada pessoa recebe uma cor diferente.



Figura 3.4. - Mapa com o percurso do olhar de um usuário.

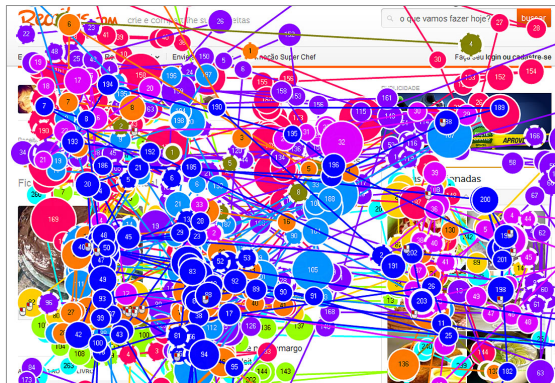


Figura 3.5. - Percurso do olhar dos usuários que visualizaram uma mesma página, onde cada pessoa é representada por uma cor diferente.

3.1.8.5.

Mapa com fixações próximas agrupadas em áreas (*cluster*)

O equipamento de *eye tracking* pode gerar um mapa que agrupa em áreas as fixações de olhares mais próximas. Na figura 3.6., é possível observar uma área em verde e outra em amarelo. Cada uma dessas áreas reúne as visualizações de um usuário na página. Essa imagem tem o propósito de mostrar um exemplo do tipo de mapa que pode ser gerado. No entanto, ela não é indicada para analisar os dados de somente um usuário. Como cada área é apresentada com uma porcentagem, que indica a quantidade de visualizações de todas as pessoas, ao utilizar os dados de apenas um participante, todas as áreas exibem o *label* 100%.

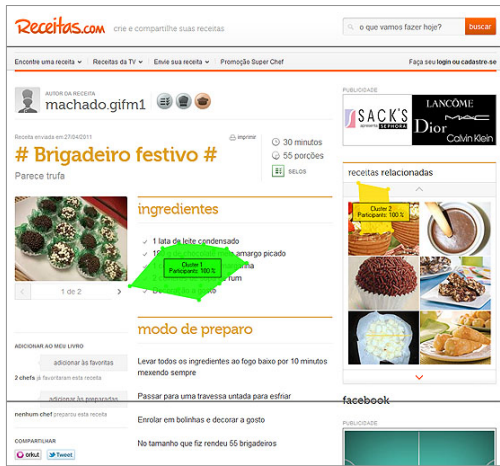


Figura 3.6. - Mapa com as fixações de olhares mais próximas de um usuário, agrupadas por áreas.

Na figura 3.7., é possível perceber a diferença em relação à figura 3.6., pois a figura 3.7. exibe os dados das visualizações de dez participantes que navegaram nessa página. Cada área apresenta sua própria porcentagem de visualização, considerando os dados de todos os participantes. Através da imagem, percebe-se que a área do título da receita foi vista por 90% a 100% dos usuários, enquanto a segunda foto do box de receitas relacionadas foi vista somente por 50% dos participantes.

Ao contrário do mapa do percurso do olhar, a representação através de cores no mapa com fixações próximas agrupadas em áreas não indica cada participante. Portanto, essas cores servem somente para diferenciar e agrupar as áreas com as fixações dos olhares de diferentes usuários.



Figura 3.7. - Mapa com fixações próximas agrupadas em áreas. Reúne todas os pontos visualizados por todos os usuários que navegaram nessa página. Diferente do mapa de percurso

de olhar, as cores não representam usuários. Nesse caso, só servem para diferenciar as áreas visualizadas.

3.1.8.6.

Análise por áreas de interesse (*area of interest* ou AOI)

O *software* que é utilizado junto com o *eye tracker* permite desenhar áreas de interesse em uma página *web*, em um vídeo ou em uma imagem que esteja sendo analisada. Dessa forma, o pesquisador pode separar áreas específicas do objeto de estudo, desenhando polígonos, onde ele pode fazer uma análise mais detalhada relativa àquela área cadastrada no sistema. Durante essa análise, é possível verificar a quantidade de fixações de olhares naquela parte, o tempo total de fixações, para qual ponto o usuário olhou antes de visualizar aquela área em particular, qual foi o último ponto visto dentro da figura, fazer comparações entre áreas, além de tantos outros tipos de análise.

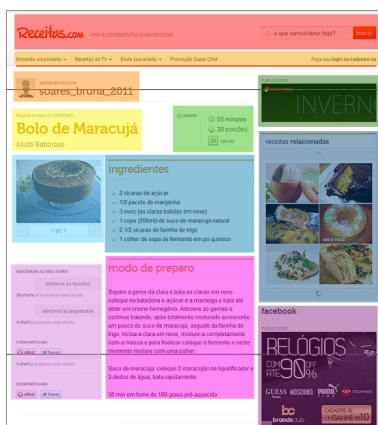


Figura 3.8. - Áreas de interesse (AOI) desenhadas na página.

3.1.8.7.

Análise do vídeo de cada usuário

Ao realizar *eye tracking* em uma página *web*, a ferramenta permite a gravação de toda a navegação do usuário, além de vincular a imagem e o áudio de uma *webcam* ao arquivo gerado. Se, ao realizar a técnica, houver uma sala de espelhos, com observadores atrás do espelho, também é possível exibir o vídeo para esses observadores. Dessa forma, as pessoas que estão na sala de observação conseguem assistir a navegação do usuário e, ao mesmo tempo, verificar para onde ele está olhando naquele instante. Na figura 3.9., a letra 'A' mostra a visão da sala de observação e a letra 'B', o percurso do olhar na tela da tv, enquanto o usuário navega por um *website*.

Recomenda-se, em qualquer aplicação de *eye tracking*, sempre assistir aos vídeos para complementar a análise dos dados, não deixando isso somente a cargo dos mapas de calor, que apresentam algumas restrições técnicas (conforme relatado anteriormente) que podem levar ao erro de interpretação do pesquisador (em *websites* com elementos dinâmicos - como *light-boxes* ou *layers* - os mapas de calor não reconhecem a diferença de visualizações entre estes elementos e os que estavam atrás deles, gerando um resultado único para elementos distintos).

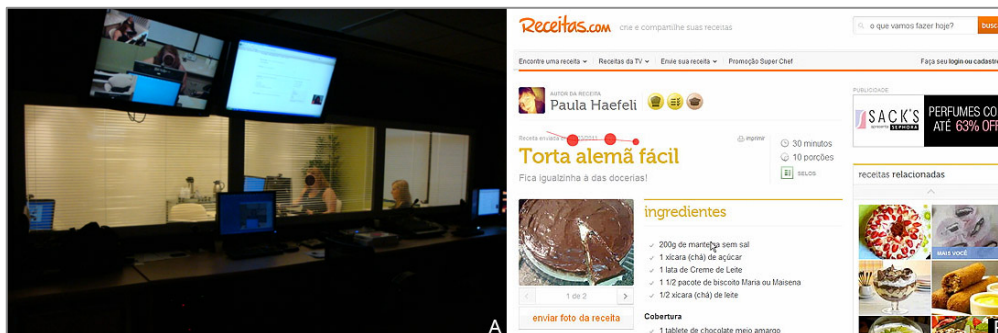


Figura 3.9. – A letra 'A' mostra a visão da sala de observação no laboratório de usabilidade da Globo.com, enquanto o usuário navega no Receitas.com e o seu percurso do olhar é exibido na tv, representado pelas bolas vermelhas na imagem 'B'.

3.1.8.8.

Análise através da reunião de fixações de olhares em um único mapa, para verificar elementos fixos

Os tipos de análise que foram apresentadas até o momento também permitem a reunião de dados a partir das visualizações de todos os usuários que acessaram determinada página. As figuras 3.10. e 3.11. exibem os diferentes tipos de mapa com as fixações dos dez os usuários que visualizaram uma página específica. Esse tipo de análise, que reúne os dados em um único mapa, é mais indicada porque permite que o pesquisador utilize os dados de toda a amostra para fazer uma consolidação das áreas mais e menos quentes da página. Desta forma, é possível fazer comparações entre o percurso do olhar dos usuários e verificar se houve, ou não, um padrão de consumo do produto. Portanto, os dados agrupados reforçam os resultados, o que reforça a recomendação pelo seu uso ao invés da análise individual.

No caso do mapa de calor, o formato que reúne todos os dados da amostra é mais adequado para análise do que o formato que reúne dados individuais. Mas no caso do percurso do olhar, fica mais complicado analisar algumas informações sobrepostas e embotadas na imagem, em função do uso

de todos os dados da amostra reunidos. Dessa forma, recomenda-se o uso dos dados individuais para analisar cada percurso e o uso dos dados agrupados para concluir a análise de forma comparativa.

As situações descritas acima referem-se ao agrupamento dos dados das visualizações de todos os usuários que acessaram a mesma página (mesma URL). Sendo assim, se apenas três, dos dez usuários, tiverem acessado uma página, o mapa com o agrupamento será referente às fixações destas três pessoas. Quando a tarefa dada ao usuário é mais exploratória (livre), pode ser que algumas páginas não sejam visitadas por todos os participantes. Talvez essas páginas nem alcancem o número mínimo seguro para uma análise de *eye tracking* (assunto que será relatado no próximo tópico deste capítulo, ou seja, o número adequado de participantes para a aplicação da técnica). Portanto, ao pensar em uma tarefa, é preciso ter em mente que esse tipo de coisa pode acontecer e, por isso, deve-se pensar muito bem sobre qual página será mostrada primeiro para os usuários, no momento em que o site for apresentado para eles. Essa página, inclusive, é a única que o pesquisador pode ter certeza de que será acessada por todas as pessoas da amostra.

Para contornar esse tipo de situação, recomenda-se um outro tipo de agrupamento, classificado como “agrupamento de dados para análise de elementos fixos em páginas do mesmo modelo”, que foi desenvolvido nesta pesquisa de mestrado para a análise dos dados do objeto de estudo, ou seja, o *site* Receitas.com. A primeira página de receita foi visualizada por todas as pessoas, pois foi a primeira a ser mostrada para os participantes. A partir dessa página, como os usuários estavam em uma navegação livre e exploratória do *site*, cada pessoa podia clicar em qualquer uma das milhares de receitas disponíveis. Portanto, com exceção da primeira página, as demais tiveram poucas pessoas as visualizando. Ao todo, foram quarenta e sete páginas de receitas acessadas por dez pessoas, sendo que algumas dessas páginas foram visualizadas somente por uma pessoa.

Como forma de ter uma análise mais segura da página de receita, aproveitando os dados das quarenta e sete páginas acessadas, criou-se outro agrupamento para analisar os componentes fixos. Portanto, através da ferramenta, foi criado um agrupamento de todas as páginas que tinham exatamente o mesmo modelo, mas que apresentavam conteúdos (texto e foto) de receitas diferentes (URLs diferentes). Então, gerou-se cada tipo de mapa com as visualizações de todas as páginas deste mesmo modelo, vistas por todos os usuários. A única variação entre essas receitas era a altura da página, pois uma

receita com mais texto aumentava o comprimento da página e vice-versa. Desta forma, a imagem da página mais comprida foi selecionada para a sobreposição dos dados, tornando possível analisar todos os elementos fixos, ou seja, que não mudavam de posição nas páginas como o *header* (cabeçalho), o menu, a área do título, a foto da receita, *banners* e box de receitas relacionadas. O texto descritivo da receita, localizado na área de altura variável da página, e o elemento de comentários, que fica abaixo deste texto (e cuja posição também varia de localização), não foram analisados nesse tipo de mapa com o agrupamento de todas as visualizações de todas as páginas do mesmo modelo (justamente por sofrerem variações de altura ou posição, ficando mais acima ou mais abaixo). No entanto, esses dois elementos puderam ser analisados a partir das visualizações dos dez usuários na primeira página de receita apresentada.

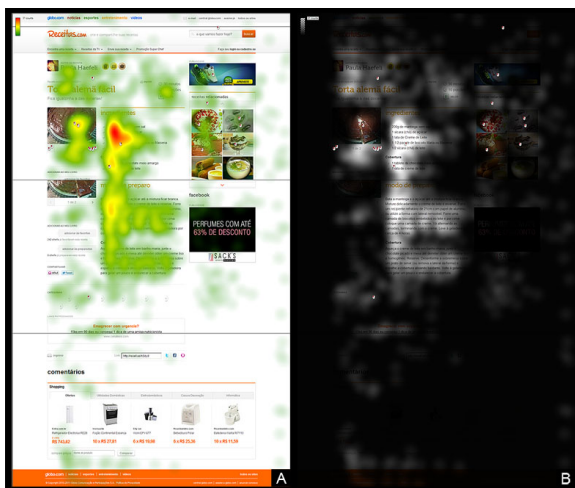


Figura 3.10. - A imagem apresenta a sobreposição dos dados das fixações dos dez usuários que acessaram essa página. A letra 'A' exibe o mapa de calor e a letra 'B' o mapa de calor invertido (ou mapa de opacidade).



Figura 3.11. - A imagem apresenta a sobreposição dos dados das fixações dos dez usuários que acessaram essa página. A letra 'C' exibe o mapa de percurso do olhar e a letra 'D' o mapa com os *clusters*.

3.1.8.9. Quantidade de usuários

Em relação ao número de participantes, NIELSEN e PERNICE (2010) recomendam um número mínimo de trinta usuários, para se considerar o resultado de um *eye tracking* como seguro. Por outro lado, há estudos na área, como o apresentado na UPA Conference 2011 (*Usability Professional Association*), onde Ania RODRIGUEZ, Marc RESNICK e Eugenio SANTIAGO mostraram suas conclusões sobre o estudo *Head to head: usability testing with and without eye-tracking*. No estudo em questão, os autores discutiram sobre a possibilidade gerar o mapa de calor das mesmas páginas, para uma quantidade diferente de usuários, de forma a comparar o resultado e verificar se as diferenças seriam significativas. Foram gerados oito mapas de calor, cada um com uma quantidade diferente de pessoas (uma pessoa, cinco pessoas, dez pessoas, dezesseis pessoas, vinte pessoas, vinte e cinco pessoas, trinta pessoas e trinta e seis pessoas). A conclusão do estudo foi de que o mapa de calor com dez pessoas é muito semelhante ao mapa com a visualização de trinta e seis pessoas, apontando uma variação muito pequena entre os dois (figura 3.12.). Após este estudo, os autores passaram a realizar outras aplicações da técnica de *eye tracking* utilizando uma amostra de dez usuários.



Figura 3.12. - Slide da palestra de Ania RODRIGUEZ, Marc RESNICK e Eugenio SANTIAGO, sobre o estudo *Head to head: usability testing with and without eye-tracking*, apresentado na UPA 2011: a figura mostra oito mapas de calor que reúnem dados das visualizações das pessoas que participaram o estudo. A primeira imagem apresenta os dados de uma pessoa, a segunda de cinco, a terceira de dez, a quarta de dezesseis, a quinta de vinte, a sexta de vinte cinco, a sétima de trinta e a oitava de trinta e seis pessoas. O estudo conclui que o resultado para dez pessoas é semelhante ao resultado para trinta e seis pessoas.

3.2.

Técnicas para projetar e testar a experiência do usuário em produtos desenvolvidos com métodos ágeis

Em 1989, Jakob NIELSEN apresentou o artigo “Engenharia de usabilidade com desconto” (*Usability engineering at a discount*, em inglês) na 3ª Conferência Internacional sobre Interação Humano-Computador, realizada em Boston. Na ocasião, foi apresentado oficialmente o conceito de *Discount Methods*, que pode ser traduzido como métodos mais baratos e mais rápidos para avaliar a usabilidade. Esse conceito englobava a ideia de testes simplificados, com poucos usuários, em conjunto com a utilização da técnica de pensar em voz alta (*think aloud*) e que uma quantidade de cinco usuários seria suficiente para descobrir resultados utilizáveis (mais que isso, os resultados acabavam se repetindo). Indicava-se fortemente o uso de cenários e da técnica de protótipo em papel, devido a sua facilidade de produção e execução, e também por permitir testes logo no início do desenvolvimento de um produto, além da sua repetição em outros momentos. O autor também indicava o uso da técnica de avaliação heurística¹⁰, para verificar se o produto estava sendo desenvolvido de acordo com as boas práticas da usabilidade. Desta forma, as premissas da abordagem de *Discount* eram:

- Todos os dados são dados;
- Qualquer coisa é melhor do que nada;
- É melhor testar com um do que com nenhum.

NIELSEN também passou a utilizar o termo *Guerrilla HCI* (interação humano-computador em guerrilha) para essa abordagem na forma de métodos mais rápidos e baratos, quando as empresas não têm tempo nem orçamento para investir em pesquisas maiores e mais custosas. Desta forma, a abordagem de guerrilha seria uma maneira das empresas começarem a implementar pesquisas, tendo o mínimo possível como ponto de partida, para depois aumentarem gradualmente o investimento, de acordo com os resultados obtidos. NIELSEN (1994) defende que empresas que utilizam os *Discounts Methods* são capazes de obter resultados significativos, passando a dar mais importância para

¹⁰ **Avaliação heurística:** um pequeno grupo de especialistas em usabilidade examina uma interface e procura por problemas. A avaliação da interface é feita de maneira isolada, evitando que os achados de um especialista sejam influenciados pelos de outro. Posteriormente, comparam-se os resultados para estabelecer o nível de gravidade dos problemas encontrados (SANTOS e MORAES, 2000).

as pesquisas durante a metodologia de projeto e, conseqüentemente, a construir seus próprios laboratórios de usabilidade.

Segundo NIELSEN e NODDER (2008), os métodos ágeis foram desenvolvidos por programadores para resolver problemas da implementação de sistemas. Se o cuidado necessário não for tomado, o design de interação e a usabilidade podem ser facilmente esquecidos, caso sejam considerados como um efeito colateral da implementação, fazendo com que este tipo de abordagem se volte contra toda a conquista da área nos últimos anos. Um time de desenvolvimento precisa utilizar métodos de design de interação e de usabilidade para construir produtos com uma boa experiência de uso. Mas, para o design ser levado a sério nesse tipo de cenário, a disciplina também precisa fazer parte das “histórias do usuário” que são a base do *sprint*. No entanto, NIELSEN e NODDER (2008) alertam que na metodologia ágil o desenvolvimento de um produto é dividido em partes menores, que são entregues uma de cada vez, fazendo com que essa abordagem ameace o conceito de uma integração total na experiência do usuário, se aproximando de uma colcha de retalhos. O ideal é fazer com que as diferentes funcionalidades que são projetadas trabalhem de forma consistente para ajudar ao usuário a interagir adequadamente com aquele produto. Mas os profissionais responsáveis por projetar a experiência do usuário que trabalham em equipes de desenvolvimento de produtos em métodos ágeis, geralmente, não têm a visão do todo, uma vez que só devem detalhar o que será desenvolvido no próximo *sprint*. Portanto, NIELSEN e NODDER (2008) indicam que a equipe de design trabalhe um ou dois *sprints* a frente da equipe de desenvolvimento, para ter tempo suficiente de projetar e testar, antes da interface ser implementada, ou seja, antes do código ser programado. Além disso, os autores também recomendam a utilização de cenários, *storyboards* e protótipos em papel, conforme relatado nos *Discount Methods*. No entanto, testes com usuários devem ser menores, porém contínuos e realizados em ciclos, além de serem testes rápidos, do tipo RITE (este conceito será relatado adiante, ainda neste capítulo).

Como forma de ajudar no planejamento inicial de um produto que começará a ser desenvolvido em método ágil, muitos autores recomendam um *sprint* diferenciado, antes de iniciar o ciclo de desenvolvimento. A ideia é que seja um tempo para a equipe se planejar e fazer uma imersão no universo do produto novo, bem como permitir que a equipe de designers possa começar a desenhar *sketches* (rascunhos da interface), cenários e fazer protótipos em papel para testar com usuários, antes do produto começar a ser construído.

Nesse momento de planejamento, a equipe de programação pode fazer estudos sobre as tecnologias disponíveis e preparar as ferramentas necessárias, assim como o ambiente, para começar o desenvolvimento a seguir. Esse tempo anterior ao início dos ciclos de desenvolvimento ágil são classificados com nomenclaturas diferentes, como *Sprint Zero*, *Discovery* ou *Spike*, embora seu propósito seja o mesmo.

PATTON (2008) defende o uso da técnica de *user story mapping*, ou “mapeamento de histórias do usuário”, como uma boa forma de tentar fazer com que o time tenha uma visão mais abrangente do projeto, sem ter que detalhar tudo nesse momento inicial do produto. A ideia é fazer com que todas as pessoas envolvidas tenham uma visão unificada do produto, assim como do possível lançamento de versões e quais funcionalidades fazem parte de cada versão. Então, deve-se fazer uma listagem de atividades e tarefas do usuário, colocando-as em uma espécie de mapa com escala de prioridade e dependência entre elas. PATTON (2008) e SY (2007) também recomendam que a equipe de experiência do usuário trabalhe alguns *sprints* a frente do time de desenvolvimento, fazendo testes rápidos e frequentes através do uso de técnicas como protótipo em papel. Ao invés do uso de termos como *Discout* e *Guerrilla*, os autores utilizam *Lightweight Methods* e *Agile Usability* ao se referirem sobre o uso destas técnicas para uma abordagem de avaliação de usabilidade ágil, ou seja, rápida, barata e simples.

COOPER, REIMANN e CRONIN (2007) descrevem um processo parecido, mas utilizam o termo *Lean UX*, ou *Lean Usability*, ao se referirem ao uso dessas técnicas recortadas e frequentes. Porém, os autores recomendam uma abordagem inicial realizada através de rodadas de entrevistas com usuários e clientes, internos e externos, garantindo uma boa forma de reunir material para iniciar a produção de *personas*, desenvolver *sketches* e fazer a modelagem do novo sistema.

A figura 3.3. descreve a proposta de NIELSEN e NODDER (2008), inspirada no trabalho apresentado por Desirée SY e Lynn MILLER, sobre a criação de um ciclo paralelo, onde o time de desenvolvimento trabalha em um *sprint* e a equipe de UX - *User eXperience* em um outro *sprint* à frente. Desta forma, o processo começa no *sprint* zero, onde é realizado o planejamento e o levantamento de dados. Este ciclo serve para criar a visão do produto que será desenvolvido. Na sequência (ciclo um), a equipe de desenvolvimento começa a implementar as partes técnicas mais complexas, além de construir o ambiente onde o produto irá funcionar. No mesmo ciclo um, a equipe de design cria o

projeto da interface que será implementada pela equipe de desenvolvimento no ciclo dois, além de fazer o levantamento de informações necessárias para o projeto do ciclo três. No ciclo dois, a equipe de desenvolvimento implementa o projeto desenvolvido pela equipe de design no ciclo um, enquanto esta já testa o que foi implementado no ciclo, em paralelo com o projeto da interface do ciclo três e o levantamento de dados do ciclo quatro, e assim sucessivamente.

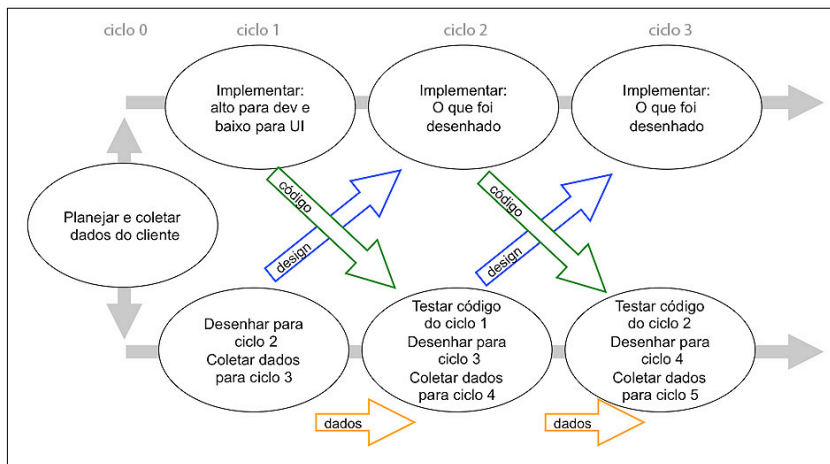


Figura 3.13. - Processo de desenvolvimento paralelo, onde a equipe de design trabalha um ciclo a frente da equipe de programação, segundo NIELSEN e NODDER (2008), com base nos estudos apresentados por Desirée SY e Lynn MILLER. Tradução e desenho de Christiane Melcher.

A grande questão é que, nos materiais disponíveis para consulta, não é possível encontrar todo o detalhamento necessário para garantir o teste de qualquer tipo de produto desenvolvido com metodologia ágil e, principalmente, como os projetos mais complexos podem ser testados, uma vez que nem sempre é possível desenvolver protótipos e manter a equipe de design trabalhando tantos *sprints* a frente, conforme o indicado. Pelo levantamento realizado com diversos times de desenvolvimento em *agile*, algumas funcionalidades podem, e devem, ser testadas de maneira rápida e frequente, talvez com o uso de protótipos em papel e outras técnicas sugeridas pelos autores da área. Mas, na prática dos grande portais de mídias digitais brasileiros, nem sempre tais técnicas serão capazes de produzir resultados significativos para as questões mais complexas a serem investigadas. No entanto, no caso deste tipo de cenário, não encontra-se outras recomendações, além das supracitadas (e que ainda serão detalhadas abaixo). Por mais que haja uma vontade de permitir que os profissionais de UX - *User eXperience* trabalhem em *sprints* à frente, muitas vezes isso não acontece como a teoria do mundo ideal determina.

Apresenta-se a seguir as técnicas recomendadas para o projeto e teste de produtos desenvolvidos em métodos ágeis e, no último tópico, apresenta-se uma discussão sobre algumas das recomendações descritas neste capítulo, que foram utilizadas como base para o início desta pesquisa de mestrado.

3.2.1.

O projeto da experiência de uso: *Personas*, Cenários e *Sketches*

A técnica de *personas* foi proposta oficialmente em 1999 e se popularizou através do livro *The inmates are running the asylum: why high tech products drive us crazy and how to restore the sanity*, de Alan Cooper, que pode ser traduzido como “Os pacientes estão controlando o asilo: porque produtos de alta tecnologia nos levam à loucura e como restaurar a sanidade”. Essa técnica apresenta uma forma de agrupar características importantes, de diferentes tipos de usuários de um produto, em uma espécie de perfil fictício (classificado como *persona*) que engloba características importantes de grupos de usuários reais e fornece os dados necessários para que os desenvolvedores de sistemas digitais possam projetar a interface de acordo com esse perfil fictício. Portanto, os grupos de usuários do produto têm suas necessidades representadas através destas *personas*. O livro teve sua segunda edição lançada em 2004. COOPER (2004) cita os seguintes benefícios a partir da utilização da técnica de *personas*:

- Ajuda o time a compartilhar um entendimento específico e consistente sobre vários grupos de usuários. Os dados sobre estes grupos podem ser colocados em um contexto apropriado, de forma a permitir uma maior compreensão em diferentes cenários;
- As soluções propostas podem ser guiadas por quanto tempo elas forem capazes de satisfazer as necessidades das *personas*. Funcionalidades podem ser priorizadas com base em quanto tempo elas forem capazes de atender as necessidades de uma ou mais *personas*;
- Fornece uma “face humana”, de modo a permitir a concentração de esforços para atender as necessidades das *personas* que representam a distribuição demográfica de consumidores daquele produto.

PATTON (2008) ressalta que os designers podem obter informações sobre usuários para permitir o desenvolvimento das *personas* através de entrevistas, observações, questionários, grupos de foco e etc. O autor também indica a priorização de alguns tipos de usuário e torná-los alvo. Em um sistema padrão, pode-se considerar dois ou três tipos de usuário como representantes mais

importantes. Para que isso seja possível, recomenda-se que sejam criadas fichas com os perfis dos usuários mais importantes, para ajudar na compreensão de quais características devem ser consideradas na criação posterior das *personas*. Nestes perfis, deve-se listar o número de usuários que ocupam esse perfil, responsabilidades gerais, atividades, habilidades com informática, metas de como o sistema pode ajudar esse usuário a alcançar seus objetivos, quais são os contextos de uso do *software* que será projetado, quais são os problemas relatados com o uso de *softwares* similares e qual a frequência de uso. A partir disso, cria-se as *personas* através da reunião das características mais importantes de cada tipo de usuário. Uma boa ficha de persona pode conter seu nome, cargo no trabalho, alguma citação, dados demográficos relevantes, uma imagem que a represente, descrições de seus objetivos, motivações, dificuldades encontradas, relação com aquele tipo de produto, atividades e quais tipos de usuário representa.

Uma vez que as *personas* de um projeto estão criadas, o próximo passo é criar cenários de uso para o produto, envolvendo cada *persona* e suas necessidades. Em seguida, pode-se criar as histórias do usuário, que são a representação da funcionalidade que será desenvolvida em algum momento nos próximos *sprints*. Então, uma vez que *personas*, cenários e histórias estão criados, passa-se detalhar as telas com a utilização de papel e caneta, através do desenho de cada cenário de uso, que pode ser considerado como um esboço (*sketch*). Alguns autores também recomendam a criação de *storyboards* para descrever cada interação do usuário com o sistema.

Conforme relatado anteriormente, os autores recomendam que os profissionais de UX - *User eXperience* trabalhem à frente do time de desenvolvimento. Sendo assim, através dos ciclos de projeto, esses profissionais devem criar os cenários, desenvolver os *sketches* e testar, seja com prototipagem em papel ou através de testes de usabilidade mais ágeis, como com a aplicação do método RITE, que será apresentado adiante, ainda neste capítulo.

3.2.2.

Teste de usabilidade com protótipo: alta ou baixa fidelidade

Segundo GOODWIN (2009), um protótipo pode ser de alta-fidelidade (bem fiel à experiência final que o usuário terá ao interagir com aquele produto) ou de baixa-fidelidade (apenas algo que se aproxime dessa experiência, mas que não

seja exatamente igual). De acordo com a autora, um modelo de protótipo de alta-fidelidade bastante utilizado, costuma ser construído com base no acabamento mas raramente na reprodução de funcionalidades e mecânicas do produto final. Um protótipo de *software* de alta-fidelidade também deve ser clicável, apresentar mecanismos realistas de entrada e saída de dados, além de permitir que o comportamento de uso seja próximo do real, com cliques e *feedbacks* de botões, por exemplo. GOODWIN (2009) também acrescenta que um protótipo de alta-fidelidade é importante para testar interações (tanto as complexas quanto as sutis) através de um *display* que tenha manipulação direta. Porém, apesar de ser de alta-fidelidade, nunca será o suficiente para se parecer com um sistema real, embora os usuários tenham a expectativa de que o botão responda de forma tão sensível como na versão final do *software* que será lançado.

Para GOODWIN (2009), os protótipos de baixa-fidelidade, como os *sketches* (esboços) em papel, são rápidos e baratos de produzir, além de não criarem expectativas irreais nos usuários. Infelizmente, esses protótipos podem introduzir um certo tipo de erro, simplesmente pela falta de refinamento. A autora cita, por exemplo, um protótipo que não possui um visual refinado, que provavelmente não passará a experiência adequada sobre a hierarquia da informação, o que tornaria mais difícil para um participante escolher o que é importante na página. Além disso, convenções típicas de um *wireframe*, como o uso de um retângulo com um “X” para representar uma imagem, muitas vezes podem não fazer sentido para os usuários. Na representação de uma tela de formulário, por exemplo, esses detalhes podem não ser tão críticos. Mas para interfaces mais complexas, esses detalhes podem ser essenciais.

GOODWIN (2009) recomenda o aproveitamento dos desenhos criados para a apresentação dos diferentes cenários de uso, utilizando-os também para a realização do protótipo em papel, permitindo uma maior aproximação, em termos de clareza e hierarquia da informação, com o sistema que será desenvolvido. Como a indicação é sempre desenhar *personas* e cenários, estes desenhos também podem servir para o protótipo de baixa-fidelidade.

Segundo SYNDER apud SANTA ROSA e MORAES (2008), a prototipagem em papel é uma variante do teste de usabilidade, na qual os usuários representativos do público-alvo do sistema executam tarefas reais, interagindo com uma versão em papel da interface, manipulada por um especialista em usabilidade que se comporta como se fosse o próprio computador, sem explicar como a interface pretende funcionar. NIELSEN apud SANTA ROSA e MORAES (2008) reforça que a prototipagem em papel é uma

técnica interativa, rápida e barata, que oferece um dos melhores e mais rentáveis métodos de obter uma visão de design nas fases iniciais do projeto. De acordo com GOMES (2007), essa técnica deve ser encorajada para a exploração de conteúdo e estrutura, enquanto que os protótipos de alta-fidelidade devem servir para vender ideias à pessoas, bem como testar questões mais técnicas. Para KLEE (2000) apud SANTA ROSA e MORAES (2008), a prototipagem em papel deve ser usada com o objetivo de medir a função e o fluxo de interação da interface e não o aspecto visual. Porém, como desvantagem, a técnica engloba a dificuldade de simular o comportamento de alguns elementos da interface, como *scrollbars*, a transmissão de informação através de cores ou animações e o fato deste tipo de metodologia não permitir a detecção de todos os tipos de problemas de usabilidade.

Esta pesquisa de mestrado mantém seu recorte, justamente, em produtos que não têm espaço no seu planejamento de desenvolvimento ágil para realizar testes de usabilidade através de protótipos, ou que, mesmo havendo tempo para isso, não é possível simular como será a experiência devido à complexidade exigida para a prototipação. Portanto, a pesquisa defende a premissa que algumas questões podem ser investigadas de forma ágil e recortada, mesmo quando uma funcionalidade ainda está sendo projetada. Porém, quando a equipe não tem tempo para criar protótipo, o teste só poderá ser realizado em um momento posterior, fazendo com que as questões a serem esclarecidas se acumulem, necessitando de uma metodologia para investigar a maior quantidade de questões no menor tempo possível, sem perder a confiabilidade dos resultados obtidos.

3.2.3.

Método de avaliação para testes rápidos e iterativos (RITE)

RITE significa *Rapid Iterative Testing and Evaluation method*, que pode ser traduzido como “método de avaliação para testes rápidos e iterativos”. Trata-se de uma abordagem voltada para testes rápidos em produtos. NIELSEN e NODDER (2008) indicam a técnica como adequada para testar a usabilidade de produtos desenvolvidos com métodos ágeis.

MEDLOCK et al (2002) ressaltam que a técnica se difere do teste tradicional de usabilidade, por dar grande ênfase à extrema rapidez nas mudanças que são feitas no produto testado, além da verificação da eficácia dessas mudanças. Segundo os autores, alterações na interface devem ser feitas

o quanto antes possível, sempre que um problema for identificado e haja uma solução clara para resolvê-lo. Caso um problema seja identificado após um usuário interagir com a interface, a equipe pode decidir por realizar as alterações necessárias no protótipo, mesmo antes de fazer outro teste com um segundo usuário. Esse processo pode ser repetido até o time considerar que foi aplicado de forma suficiente. As questões identificadas pelos usuários que participam dos testes são classificadas em quatro categorias:

1. Questões que parecem ter uma causa óbvia, assim como uma solução clara, que pode ser implementada com rapidez;
2. Questões que parecem ter uma causa óbvia, com uma solução também óbvia, mas que não podem ser implementadas tão rápido, ou dentro do prazo do teste atual;
3. Questões que parecem não ter uma causa tão óbvia, consequentemente, cuja solução também não é tão clara;
4. Problemas que podem ser causados por outros fatores, como pelo *script* de teste ou pela interação do participante, etc.

Para os problemas da categoria 1, recomenda-se implementar a solução e testar com o participante seguinte. Para os problemas da categoria 2, deve-se iniciar a implementação da solução e testar, quando for possível, no protótipo remodelado. Para os problemas das categoria 3 e 4, deve-se coletar mais dados com os usuários, para verificar se esses problemas mudam de classificação (para as categorias 1 ou 2), ou se são reclassificadas como questões “não categorizáveis”.

De acordo com MEDLOCK et al (2002), deve haver um acordo prévio, antes do teste, sobre quais tarefas todo usuário do sistema deve, obrigatoriamente, ser capaz de desempenhar, sem exceções. Esse acordo é importante para enfatizar a importância das questões observadas no teste. Também é importante existir a possibilidade de realizar as mudanças necessárias de forma bem ágil, em menos de duas horas, para permitir que o teste seja aplicado com um segundo usuário, ou pelo menos antes do próximo dia de teste agendado. Isso significa que o tempo de desenvolvimento e os recursos devem ser reservados somente para realizar tais ajustes. Além disso, após cada teste, ou cada dia de sessões de avaliações, um tempo deve ser destinado a uma conversa da equipe, para analisar as descobertas e tomar as providências, caso algum ajuste seja necessário. Após cada mudança, é preciso testar novamente, para garantir que esses ajustes foram adequados e não causaram novos problemas. Mas, apesar dos autores relatarem que não há uma

quantidade mínima de usuários definida para realizar os testes, eles também recomendam que a tabela de LEWIS (1991) seja seguida, para calcular a distribuição binominal.

Tamanho da Amostra	Probabilidade de Descoberta do Problema	Porcentagem de Problemas Descobertos
0	0.000	0.0
1	0.500	50.0
2	0.750	75.0
3	0.875	87.5
4	0.938	93.8
...		
Infinito	1.000	100.0

Tabela 1. - Distribuição de Probabilidade Binominal: retornos decrescentes. LEWIS (1991)

No entanto, MEDLOCK et al (2002) afirmam que para essa técnica entregar bons resultados, é preciso seguir algumas regras:

- O engenheiro de usabilidade deve ter bastante experiência, tanto a respeito do universo do produto testado quanto aos problemas que os usuários costumam enfrentar ao lidar com esse tipo de produto. Sem essa experiência, é difícil prever se um problema descoberto por apenas uma pessoa também pode ser um problema para as outras pessoas.
- Se o profissional de UX - *User eXperience* não tiver experiência com o produto testado, talvez um teste de usabilidade tradicional seja mais indicado.
- Os líderes da equipe de desenvolvimento devem participar dos testes, para poder delegar os ajustes de forma rápida. Portanto, não se trata apenas de alguém que, simplesmente, vai assistir aos testes. Ao contrário disso, trata-se de pessoas realmente envolvidas e comprometidas em agir.
- O time de designers e o engenheiro de usabilidade devem ser capazes de interpretar os resultados de forma rápida, para que as decisões sejam igualmente ágeis e eficazes na hora de resolver o que será alterado no protótipo.

3.2.4. Pontos para discussão

Apesar de líderes em testes de usabilidade da Microsoft terem recomendado o uso da técnica RITE no congresso da Associação dos Profissionais de Usabilidade em 2002, além de autores renomados também defenderem seu uso, como o relatório de Jakob NIELSEN (2008) sobre

desenvolvimento ágil e *user experience*, acredita-se que ainda existem algumas lacunas não esclarecidas, não somente em relação à técnica RITE, mas também em relação às outras recomendações para a aplicação de testes de usabilidade em produtos desenvolvidos com métodos ágeis. Vários autores renomados da área, afirmam que suas recomendações são a única forma indicada para testar produtos neste cenário. No entanto, é necessário discutir algumas considerações sobre o assunto:

1. Existe uma recomendação para testar o produto com a quantidade necessária de participantes, conforme a tabela de LEWIS, onde com quatro usuários pode-se achar noventa e três por cento dos problemas. No entanto, para encontrar cem por cento dos problemas, o número de pessoas necessárias tende ao infinito. Assim, isso pode ser interpretado como algo que requer um certo tempo de análise da amostra, criando uma certa distância da agilidade pretendida com a apresentação da técnica, além de um distanciamento em relação às recomendações de outros autores, que indicam testar o produto com três ou quatro usuários. Alguns autores, inclusive, relatam que é possível testar somente com um participante, conforme as técnicas *Guerrilla* ou *Lightweight methods*, pois isso já seria melhor do que não testar o produto com nenhuma pessoa.
2. De acordo com as últimas considerações a respeito da experiência que o engenheiro de usabilidade deve ter, presume-se que os profissionais de UX - *User eXperience* alocados no desenvolvimento devem ser *seniores*, não só em relação ao conhecimento de usabilidade, mas também em relação ao domínio do produto estudado. Desta forma, esses profissionais são capazes de prever se um problema encontrado pode se repetir para outras pessoas, ou se é algo isolado para aquele usuário específico que estava testando o produto. Mas, vários autores afirmam que o profissional de UX - *User eXperience* não deve projetar somente de acordo com as suas experiências ou crenças, tendo o dever de verificar como os usuários interagem com aquele produto. Portanto, questiona-se a abordagem de fazer ajustes logo depois de testar a interface com apenas uma pessoa. Na verdade, acredita-se que essa abordagem seja precipitada. Também considera-se inadequado deixar essa decisão somente em função da experiência dos projetistas. Talvez, para problemas mais pontuais, tal abordagem realmente seja melhor e mais rápida. Porém, a proponente da pesquisa questiona se essa abordagem é capaz de funcionar quando os problemas são mais complexos, ou

quando não se pode desenvolver protótipos com alto grau de fidelidade em relação ao comportamento que determinada funcionalidade terá no momento do seu lançamento. Há muitas variáveis a serem consideradas nesse contexto de desenvolvimento ágil. Portanto, indicar somente um tipo de abordagem para todos os projetos desenvolvidos com esse tipo de metodologia não é adequado.

3. Uma outra questão que se deve levar em conta é o envolvimento do time de desenvolvimento. Enquanto autores como SY (2007) sugerem que o time de UX - *User eXperience* trabalhe à frente do *sprint* de desenvolvimento e, com isso, seja capaz de realizar testes, outros autores recomendam o trabalho em conjunto com os demais membros da equipe na hora de realizar os testes, como no caso do RITE. Dessa forma, a sessão de teste deve ser uma história que faz parte do *sprint* de desenvolvimento do produto. Portanto, cada vez que algo for testado, deve-se considerar o tempo de desenvolvimento do protótipo e o tempo de realização de ajustes, consequentemente, parando o desenvolvimento do produto durante aquele período. Acredita-se que essa abordagem não é a mais indicada para todos os cenários. A proponente da pesquisa também questiona se os profissionais de UX - *User eXperience*, que desenvolvem a solução, deveriam ser os mesmos que aplicam os testes com os usuários. Na verdade, acredita-se que deveria haver um distanciamento maior de quem aplica os testes, para não correr o risco de introduzir algum tipo de viés nos resultados.
4. Autores como SY (2008), NIELSEN (2007), entre outros, indicam que é possível utilizar funcionários da própria empresa, com perfil próximo ao do usuário final, para executar os testes menores e preliminares, de forma a otimizar todo o tempo de recrutamento externo. Sendo assim, usuários reais só devem ser recrutados no final da etapa de desenvolvimento de uma versão do produto. Porém, acredita-se que tal recomendação não pode ser executada por qualquer equipe de qualquer empresa. Dependendo do produto, existe a possibilidade dos funcionários não servirem como testadores preliminares, devido a um grande distanciamento entre os seus perfis e os perfis do público final do produto. A proponente da pesquisa não questiona a recomendação, mas a generalização de que há somente um tipo de abordagem para testar a usabilidade de produtos desenvolvidos com métodos ágeis.