



INTERAÇÃO HUMANO- COMPUTADOR

AULA 5

CONVERSA INICIAL

Seja bem-vindo(a) a esta aula! Nela, será apresentada a **avaliação**, um dos principais processos de *Design* de Sistemas Interativos. O objetivo desta aula é delimitar o processo de avaliação e reconhecer que não se trata apenas da avaliação de um produto acabado, mas, sim, de todo o processo de *design*.

Os designers precisam pensar muito sobre quais as características do sistema que querem avaliar e o porquê. Também é importante considerar a relação entre a forma de antecipação e a natureza da avaliação.

Você também conhecerá os processos de revisão, experimentação ou teste de uma ideia de *design*, de um *software*, de um produto ou serviço, atendendo a determinados critérios, que são apresentados nos 5 (cinco) temas:

- Apresentar a ideia geral de **avaliação**
- A avaliação por usuários experientes, especialistas e peritos
- Analisar a avaliação baseada no participante
- Apresentar a avaliação na prática
- Concluir com a apresentação de outras questões sobre a avaliação

Aproveitem e bons estudos!

TEMA 1 – A AVALIAÇÃO

A avaliação é o quarto dos principais processos do *Design* de Sistemas Interativos. Por avaliar queremos dizer **revisar**, **experimentar** ou **testar** uma ideia de *design*, um *software*, um produto ou serviço e descobrir se ele atende a alguns critérios.

O *designer* preocupa-se não apenas com as características superficiais (como o significado dos ícones, por exemplo), mas também se preocupa com o fato de o sistema ser adequado ao seu propósito, agradável, envolvente etc.

As técnicas apresentadas nesta aula permitirão ao aluno avaliar muitos tipos de produtos, sistemas ou serviços! A avaliação de diferentes tipos de sistemas ou a avaliação em diferentes contextos pode trazer determinados desafios e está intimamente ligada a outras atividades do *Design* de Sistemas Interativos, tais como:

- Entendimento
- *Design*
- Antecipação

A avaliação tem também uma dependência crítica da forma de antecipação usada para representar o sistema. Confirmam, abaixo, os pontos-chave.

Os *designers* devem se concentrar nas características de um sistema ou produto que querem avaliar. Eles precisam pensar bem sobre o estado no qual o sistema ou produto está e se é possível avaliar essas características.

- Existem métodos de avaliação baseados em especialistas.
- Existem métodos de avaliação baseados em participantes.
- Os *designers* precisam criar uma avaliação que atenda às necessidades particulares dos contextos de uso e das atividades nas quais as pessoas estão engajadas.

1.1 DESIGN CENTRADO NO HUMANO – AVALIAÇÃO

Na abordagem de *Design* Centrado no Humano, avalia-se o *design* desde as primeiras ideias. Existem **dois** tipos principais de avaliação:

Uma é feita com um especialista em usabilidade (ou *designer* de interação) para revisar alguma forma de versão antecipada de um *design*: são os **métodos baseados em especialista**.

Outra implica recrutar pessoas para usar uma versão antecipada do sistema: **são os métodos com participantes**. Sempre que possível, as pessoas devem ser representantes daquelas para quem o sistema se destina (às vezes, chamadas de “usuários finais”). Essa é a escolha preferencial para sistemas *in-house*, nos quais o *designer* tem acesso à população-alvo.

Alternativamente, os participantes podem ser outras pessoas (talvez outro *designer*, estudantes ou quem estiver por perto) convidadas para desempenhar o papel das pessoas que usarão o sistema.

As características da população-alvo podem ser captadas por meio de **personas**.

A **avaliação** acontece ao longo do processo de *design* de interação. Em diferentes estágios, diferentes métodos serão mais ou menos eficazes. A forma de antecipação dos sistemas futuros também é crítica quanto ao que pode ser avaliado. Você talvez queira avaliar os conceitos iniciais, principalmente se a aplicação for nova. Nesse caso, protótipos rápidos de papel ou mesmos *softwares* podem ajudar, se eles puderem ser produzidos rapidamente.

Avaliações de produtos concorrentes ou versões anteriores da tecnologia podem também alimentar o processo de *design* neste estágio. Durante o desenvolvimento, os *designers* têm de decidir entre opções, por exemplo, de entrada por voz e interação *touchscreen* para uma agenda de parede doméstica, ou de diferentes sequências para a ordem de processamento de funções.

1.2 PROBLEMAS DE USABILIDADE

Os testes identificarão possíveis problemas, uma vez que uma versão estável da tecnologia esteja disponível. Um **protótipo horizontal** tem de responder quando um participante ativar uma função, mas isso não requer que o sistema todo esteja completamente operacional.

Alternativamente, em um **protótipo vertical**, o sistema pode estar funcional somente em algumas partes. O importante é que ainda se tenha tempo de consertar os problemas!

1.3 AVALIAÇÃO FORMATIVA E AVALIAÇÃO CONCLUSIVA

Na avaliação formativa, o que acontece com frequência é que se pede para verificar se a interação é "amigável ao usuário" pouco antes do desenvolvimento ser completado. Muitas vezes, tudo o que pode ser alterado são pequenos detalhes, tais como a posição, a cor, ou a identificação dos botões na tela.

Se você anotar os problemas que poderiam ter sido facilmente resolvidos se identificados antes, poderá explorar esses exemplos (com muito tato), para justificar o trabalho de avaliação em estágio mais precoce no próximo projeto.

A avaliação conclusiva avalia a usabilidade de um produto **acabado**. Isso pode ser feito com base em diretrizes próprias, em padrões formais de usabilidade – tais como a ISO 9241 – ou

fornecendo as provas de usabilidade exigidas por um cliente, como o tempo para completar determinado conjunto de operações.

Departamentos do governo e outros órgãos oficiais requerem que os fornecedores atendam a padrões de acessibilidade e à legislação de saúde e segurança. No projeto citado como exemplo, os empregadores do pessoal que seria treinado no ambiente virtual exigiram provas de que tal treinamento seria tão eficaz quanto seu equivalente na sala de aula.

No *design* com abordagem participativa, os *stakeholders* ajudam os *designers* a estabelecer metas para o trabalho de avaliação. Envolver *stakeholders* traz grandes vantagens em termos de compreensão futura e uso da tecnologia. Isso, é claro, aplica-se somente a tecnologias feitas sob medida para comunidades predefinidas, mas não se aplica a produtos “de prateleira”.

TEMA 2 – AVALIAÇÃO POR USUÁRIOS EXPERIENTES, ESPECIALISTAS E PERITOS

Um método simples, relativamente rápido e eficaz de avaliação, é conseguir que um especialista em *design* de interação ou usabilidade analise o sistema e tente utilizá-lo. Isso não substitui a avaliação com o uso do seu *design* pelo verdadeiro público, mas a avaliação por um especialista é particularmente eficaz nos estágios iniciais do processo.

Especialistas logo detectam problemas comuns baseados em experiências que identificam fatores, os quais podem, de outra forma, interferir em uma avaliação por não especialistas. Adotar determinada abordagem ajuda a concentrar a crítica do especialista nos aspectos mais relevantes para o objetivo.

2.1 AVALIAÇÃO HEURÍSTICA

Aqui, os cenários usados no processo de *design* são muito úteis. A avaliação heurística refere-se a vários métodos por meio dos quais uma pessoa treinada em IHC e *Design* de Interação examina o *design* proposto para avaliar como ele se qualifica diante de uma lista de princípios, diretrizes ou heurística para o bom *design*.

Essa revisão pode ser uma discussão rápida em uma conversa com o colega ou pode ser um processo formal e cuidadosamente documentado. Existem vários conjuntos de heurística entre os

quais escolher, tanto para uso geral quanto para domínios de aplicações específicas, por exemplo, a heurística para o *design* de *Internet*.

Conheça alguns princípios – ou heurísticas – do *design*:

Figura 1 – Princípios ou heurísticas do *design*



Fonte: André Roberto Guerra.

O ideal é que várias pessoas com experiência no *Design* de Sistemas Interativos avaliem a interface. Cada especialista anota os problemas e a heurística relevantes, sugerindo soluções onde for possível.

Também é útil que haja uma avaliação de gravidade, em uma escala de 1 a 3, por exemplo, conforme o impacto provável do problema e conforme recomendado. No entanto, observa-se o nível desapontador de correlação entre os especialistas na avaliação dos problemas de gravidade.

Os avaliadores trabalham independentemente e depois combinam resultados. Eles podem ter de estudar o material de treinamento que houver e receberem instruções da equipe de *design* sobre a funcionalidade.

Contudo, a avaliação heurística não é uma solução completa. A técnica deve ser usada com a consideração cuidadosa de pessoas e suas habilidades cotidianas e é necessária a avaliação de participantes para se chegar a um quadro realista do sucesso de um sistema.

A avaliação heurística, portanto, é uma avaliação valiosa enquanto formativa, para ajudar o *designer* a melhorar a interação em um estágio inicial. Entretanto, ela não deve ser usada como

avaliação conclusiva, para se fazer reclamações de usabilidade e outras características do produto acabado.

Se é isso o que precisamos fazer, então, devemos realizar experimentos adequados, controlados e com um número muito maior de participantes. No entanto, quanto mais controlada a situação de teste, se torna menos provável que ela se pareça com o mundo real, nos levando à questão da **validade ecológica**.

Na vida real, as pessoas realizam múltiplas tarefas, usam várias aplicações em paralelo ou em sucessão rápida, são interrompidas, improvisam, pedem ajuda a outras pessoas, usam aplicações intermitentemente e adaptam tecnologias para propósitos que os *designers* não imaginaram.

Temos estratégias imprevisíveis e complexas (mas geralmente eficazes) para lidar com o dia a dia e a sua tecnologia de apoio. As pequenas tarefas que são o foco da maioria das avaliações geralmente fazem parte de sequências longas dirigidas a objetivos que mudam de acordo com as circunstâncias.

Tudo isso é extremamente difícil de reproduzir em testes e, frequentemente, os testes são deliberadamente excluídos. Portanto, os resultados da maioria dos testes de usuário são apenas indicadores de questões de uso na vida real. Técnicos e pesquisadores estão conscientes desse problema e uma série de soluções já foram propostas. Elas incluem:

- Observações etnograficamente informadas de tecnologias de uso a longo prazo (embora isso seja realizado com maior frequência mais cedo, no ciclo de avaliação do *design*).
- Pedir aos usuários que mantenham diários audiovisuais ou escritos.
- Coligir relatos de "*bugs*": frequentemente são problemas de usabilidade.
- Consultas a centrais de ajuda.

Em um teste de avaliação heurística em larga escala, usado como exemplo, os avaliadores foram treinados para usar a técnica e depois avaliaram a interface de um editor de desenho. O editor foi, em seguida, experimentado por consumidores.

Uma comparação das constatações mostrou que muitas das questões identificadas pelos especialistas não foram sentidas pelas pessoas (falsos positivos), enquanto algumas dificuldades graves não foram vistas pela inspeção pelos padrões heurísticos.

A menos que não haja alternativa, você não deve avaliar os seus próprios *designs*! É extremamente difícil ignorar o seu conhecimento de como o sistema funciona, o significado dos ícones ou nomes no menu e assim por diante. E é provável que você dê ao *design* o benefício da dúvida ou encontre falhas obscuras que a maioria dos usuários jamais perceberá.

Quanto a deixar passar problemas, isso tende a ser o resultado de uma série de erros e concepções errôneas, frequentemente relacionadas a um conjunto de itens interligados – e não a mal-entendidos isolados.

Algumas vezes, a heurística é mal aplicada ou aparentemente acrescentada como algo secundário. Em síntese, ela apresenta poucas vantagens para a avaliação de um especialista e que os resultados de a aplicar podem ser contraproducentes. Alguns autores sugerem que técnicas teoricamente mais informadas como um acompanhamento cognitivo oferecem um suporte mais robusto para a identificação de problemas.

2.2 ENGENHARIA ECONÔMICA DE USABILIDADE

Essa abordagem de avaliação teve como pioneiro Jakob Nielsen (1993) e foi entusiasticamente seguida por muitos profissionais de avaliação pressionados pela falta de tempo. Hoje, ela é usada para qualquer abordagem de avaliação apressada na qual o objetivo é obter um retorno informado e útil o mais rápido possível.

Aqui vários especialistas em usabilidade também fazem uma avaliação rápida dos cenários concretos, de preferência acompanhados de personas e inspecionam o *design* quanto a dificuldades. A lista dos princípios de *design* acima pode ser resumida por três princípios abrangentes de usabilidade, que são:

Tabela 1 – Lista dos princípios de usabilidade no *design*

Aprendabilidade	1. Visibilidade 2. Consistência 3. Familiaridade 4. <i>Affordance</i>
Efetividade	5. Navegação 6. Controle

	7. Retorno 8. Recuperação 9. Restrições
Acomodação	10. Flexibilidade 11. Estilo 12. Sociabilidade

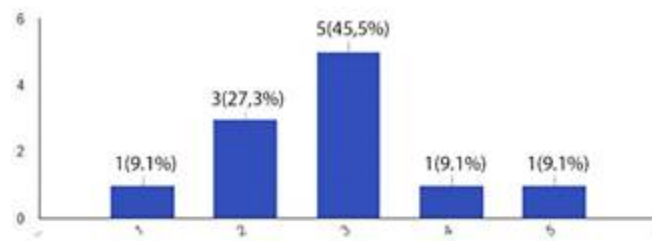
Fonte: André Roberto Guerra.

Existem diversas ferramentas disponíveis que podem ser utilizadas para a elaboração de questionários de avaliação de usabilidade, entre elas as populares ferramentas *web*, gratuitas, (Ex.: Google forms, Microsoft Teams) que serão apresentadas na aula prática, confirmam!

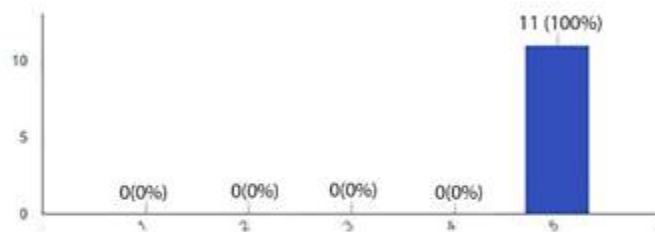
Alguns modelos de questionários são sugeridos por (Preece, 2015) e, na figura a seguir (adaptada de Carvalho, 2016), é apresentado um exemplo de questionário de especialistas:

Resultados individuais de todas as afirmações do questionário de especialistas.

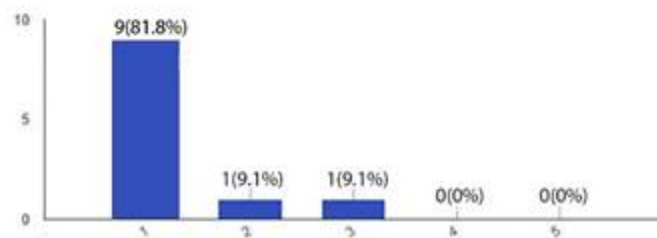
Acho que gostaria de usar este site com frequência. (11 respostas)



Achei o site fácil de usar. (11 respostas)

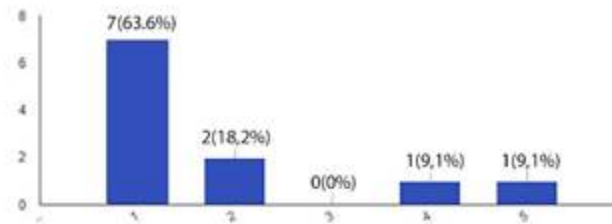


Achei este site muito inconsistente. (11 respostas)

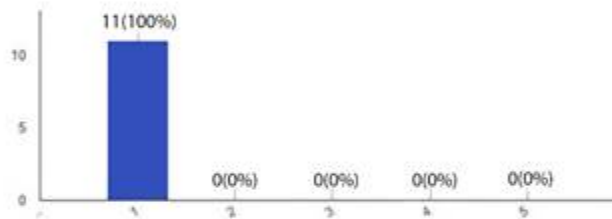


Eu me senti confortável com este site. (11 respostas)

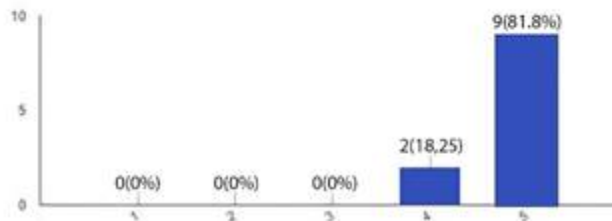
Achei o site desnecessariamente complexo. (11 respostas)



Achei que seria necessário o apoio de um técnico para poder usar este site. (11 respostas)



Imagino que a maioria das pessoas aprenderiam a usar este site rapidamente. (11 respostas)



Eu preciso aprender um monte de coisas antes de continuar usando este site. (11 respostas)

Fonte: elaborado com base em Carvalho, 2016, apêndice B, p. 58.

Se o tempo for muito curto, uma revisão rápida do *design* a partir dessa tríade pode produzir resultados razoavelmente úteis. **Acompanhamento cognitivo** é uma técnica rigorosa baseada documentos para verificar o *design* em detalhes e a lógica das etapas em uma interação. Ela é derivada da visão de cognição baseada no processamento humano da informação e está intimamente relacionada à análise de tarefa.

Em essência, o acompanhamento cognitivo consiste em um analista de usabilidade percorrendo passo a passo as tarefas cognitivas que precisam ser realizadas na interação com a tecnologia. Uma vez coletado esse material, o analista faz as quatro perguntas a seguir para cada etapa individual da interação:

- As pessoas que estão usando o sistema tentarão obter o efeito correto?
- Elas perceberão que a ação correta está disponível?
- Elas irão associar a ação correta ao efeito que estão tentando obter?

- Se a ação correta for realizada, as pessoas verão que está sendo feito progresso em direção à meta da sua atividade?

Se qualquer uma das perguntas for respondida com uma negativa, um problema de usabilidade terá sido identificado e será registrado, mas as sugestões de *redesign* não serão feitas nessa etapa!

Se o acompanhamento está sendo usado da forma como foi originalmente criado, esse processo será realizado como um exercício de grupo por analistas e *designers*. Os analistas percorrem passo a passo os cenários de uso e a equipe de *design* tenta explicar como o usuário iria identificar, realizar e monitorar a sequência correta de ações.

2.3 AVALIAÇÃO BASEADA EM ESPECIALISTA – RESUMO

Embora a avaliação baseada em especialista seja um primeiro passo razoável, ela não irá encontrar todos os problemas, particularmente aqueles que resultam de uma sequência de ações “erradas” ou que estão ligados a concepções errôneas fundamentais.

Os especialistas encontram, inclusive, problemas que, na realidade, não existem – as pessoas superam muitas das dificuldades menores usando um misto de bom-senso e experiência. Portanto, é realmente importante completar o quadro com algumas pessoas de verdade experimentando o *design* de interação.

As constatações serão sempre interessantes, frequentemente surpreendentes e ocasionalmente desconcertantes! De um ponto de vista político, é mais fácil convencer os *designers* da necessidade de mudanças se os indícios não forem simplesmente a visão de um “especialista”, particularmente se o especialista for relativamente jovem. O objetivo é testar o *design* com pessoas que representam o grupo-alvo pretendido, em condições tão próximas da realidade quanto possível.

TEMA 3 – AVALIAÇÃO BASEADA NO PARTICIPANTE

A avaliação baseada no participante tem o objetivo de envolver algumas pessoas de verdade na avaliação. Existem muitas maneiras de envolver pessoas e esse envolvimento pode se dar em vários graus de cooperação.

Podemos manter *designers* acompanhando participantes à medida que eles aprendem e lidam com um sistema ou então poderíamos deixar as pessoas sozinhas com a tecnologia e observar o que elas fazem através de um espelho de sentido único.

Um meio de maximizar os dados reunidos com uma simples sessão de teste é a utilização da **avaliação cooperativa**. A técnica é cooperativa porque os participantes não são sujeitos passivos, mas trabalham como coavaliadores. Essa técnica foi comprovada como confiável, porém, econômica em diversas aplicações.

3.1 PARTICIPAÇÃO E HEURÍSTICA

Os criadores da avaliação heurística participativa alegam que ela amplia o poder da avaliação heurística sem acrescentar muito ao esforço necessário. É sugerida a utilização de uma lista ampliada de quesitos heurísticos, mas é claro que você também poderia usar qualquer heurística.

O procedimento para o uso da avaliação heurística participativa é o mesmo da versão de especialista, mas os participantes são envolvidos como “especialistas do domínio de trabalho” ao lado dos especialistas em usabilidade e deve-se explicar a eles o que é exigido.

Codescoberta é uma técnica naturalista e informal particularmente mais adequada para captar as primeiras impressões. Ela é usada nos estágios mais tardios do *design*.

A abordagem padrão de observar pessoas interagindo individualmente com a tecnologia e possivelmente “pensando alto”, enquanto o fazem, pode variar com os participantes explorando novas tecnologias em pares.

Por exemplo, uma série de pares de pessoas pode receber um protótipo de uma nova câmera digital para que experimentem seus recursos tirando fotos uns dos outros e dos objetos na sala. Isso tende a provocar um fluxo mais natural de comentários, e as pessoas, frequentemente, irão estimular umas às outras para tentar interações e comentários que elas poderiam não ter pensado se estivessem sozinhas.

3.2 LIVING LABS

Os *Living Labs* são uma abordagem europeia para avaliação, cujo objetivo é envolver o máximo possível de pessoas na exploração de novas tecnologias. Há uma série de estruturas diferentes para

os *Living Labs*.

A ideia-chave por trás deles é a de que as pessoas estão dispostas e são capazes de contribuir para o *design* de novas tecnologias e novos serviços e que faz sentido as empresas trabalharem com essas tecnologias.

O fato de que as discussões e avaliações acontecem no contexto de vida das pessoas – e, frequentemente, com um grande número de pessoas – atribui uma forte validade ecológica aos dados.

Outra maneira de realizar uma avaliação com participantes é montar um **experimento controlado**. São apropriados quando o *designer* está interessado em determinadas características de um *design*, talvez comparando um *design* com outro para ver qual é melhor. Para fazer isso com qualquer grau de segurança, o experimento precisa ser cuidadosamente projetado e executado.

Um experimento controlado frequentemente resulta em alguns dados quantitativos – são as medidas dos valores dependentes. Esses dados podem então ser analisados com métodos estatísticos, comparando, por exemplo, o tempo médio para se fazer alguma coisa em duas condições ou o número médio de cliques.

Portanto, para realizar experimentos controlados você precisará de um conhecimento básico de teoria da probabilidade, de teoria experimental e, é claro, de Estatística.

TEMA 4 – AVALIAÇÃO NA PRÁTICA

Uma pesquisa com 103 profissionais experientes de *design* centrado no humano, conduzida no ano 2000 (Vredenburg et al, 2002), mostrou que cerca de 40% dos entrevistados utilizavam “avaliações de usabilidade”; cerca de 30% usavam “avaliação informal por especialista” e em torno de 15% usavam “avaliação heurística formal”. Esses números não dizem se as pessoas usavam mais de uma técnica.

Algun tipo de *trade-off* custo-benefício parece estar em operação. Para profissionais atarefados, a economia relativa dos métodos de revisão frequentemente compensa pela informação melhor que é obtida com os testes de usuário. Obviamente, a comunidade permanece carente de métodos que

sejam tanto econômicos no uso de recursos, quanto produtivos em termos de resultados úteis. As principais **etapas** sugeridas para avaliação são as seguintes:

- Estabeleça os objetivos da avaliação, os possíveis participantes, o contexto de uso e o estado da tecnologia; obtenha ou construa cenários ilustrando como a aplicação será usada.
- Selecione métodos de avaliação. Eles devem ser uma combinação de métodos de revisão por especialista e de métodos com participantes.
- Realize a revisão com especialistas.
- Planeje os testes com participantes; use os resultados da revisão por especialista para ajudar no enfoque.
- Recrute pessoas e organize o local e o equipamento para os testes.
- Realize a avaliação.
- Analise os resultados, documente e relate aos *designers*.

Os **objetivos** da avaliação são, em suma, decidir quais são os objetivos da avaliação ajuda a determinar o tipo de dado necessário. É útil escrever as principais perguntas que você precisa responder. A avaliação do conceito inicial no projeto Discover, por exemplo, resultou em perguntas como:

Os treinadores entendem e aceitam bem a ideia do treinamento em ambiente virtual? Eles o usariam para ampliar ou substituir os cursos de treinamento e existentes?

4.1 RELEVÂNCIA

Existe uma grande dificuldade: decidir qual o número aceitável para, digamos, a porcentagem de tarefas completadas com sucesso. Seria 95%, 80% ou 50%? Em alguns casos (raros), o cliente pode determinar esse número!

Caso contrário, um parâmetro pode ser tirado de um teste comparativo com um *design* alternativo, uma versão anterior, um produto concorrente ou a versão manual atual do processo a ser computadorizado.

Mas a equipe de avaliação ainda tem que determinar se uma métrica é relevante. Por exemplo, em um sistema complexo auxiliado por um computador, não se esperaria que a maioria das funções fossem usadas perfeitamente na primeira tentativa, não é mesmo?

E seria realmente significativo se os engenheiros que estivessem usando um *design* fossem, na média, dois segundos mais rápidos para completar um diagrama complexo em relação aos que usassem um *design* concorrente?

4.2 PESSOAS

As pessoas mais importantes em uma avaliação são as que irão usar o sistema. O trabalho de análise deve identificar as características das pessoas e representar essas características na forma de personagens.

Dados relevantes da caracterização podem incluir o conhecimento das atividades que a tecnologia deverá suportar, habilidades relativas a dispositivos de entrada e saída, experiências, educação, treinamento e capacidades físicas e cognitivas. Você deve recrutar pelo menos três e preferivelmente cinco pessoas para participar dos testes.

A amostra recomendada de três a cinco participantes é, há mais de uma década, aceita por opinião geral na prática da usabilidade. No entanto, alguns profissionais e pesquisadores alertam que pode ser pouco. Consideramos que, em muitas situações da vida real, obter até mesmo de três a cinco pessoas é difícil, de forma que continuamos a recomendar um pequeno número de participantes para os testes como parte de uma estratégia pragmática de avaliação.

No entanto, testar um número tão pequeno só faz sentido se o seu *design* se destina a um grupo relativamente homogêneo. Se o conjunto de clientes ao qual seu *design* se destina é heterogêneo, você precisará testar de três a cinco pessoas **de cada grupo**. Se o seu produto será demonstrado pelo departamento de vendas e *marketing*, vale a pena envolvê-los no teste!

Encontrar participantes representativos deve ser simples se você estiver desenvolvendo uma aplicação *in-house*. Caso contrário, os participantes podem ser encontrados nos grupos de interesse ou através de anúncios.

Estudantes frequentemente estão prontamente disponíveis, mas lembre-se de que eles são representantes de apenas um segmento da população! Se você dispõe de recursos, a remuneração pode ajudar no recrutamento. Inevitavelmente, a sua mostra irá tender a favor de pessoas cooperativas com algum tipo de interesse em tecnologia, portanto, tenha isso em mente quando interpretar os resultados.

4.3 OUTRAS OPÇÕES

Se não conseguir recrutar quaisquer participantes genuínos – pessoas que são realmente representativas dos clientes alvos – e você é o *designer* do *software*, ao menos peça para que outra pessoa tente usá-lo. Ela pode ser um dos seus colegas, um amigo, sua mãe ou qualquer pessoa em quem você confie para lhe dar uma reação brutalmente honesta. É quase certo que eles encontrarão algumas falhas de *design*.

Você terá, no entanto, que ser extremamente cuidadoso quanto até que ponto generalizar a partir das suas constatações! Por fim, considere seu próprio papel e o de outros na equipe de avaliação, se você tiver uma. Você precisará montar os testes e coletar os dados, mas até que ponto irá se envolver?

O método recomendado para os testes básicos requer que um avaliador acompanhe cada usuário e envolva-se com ele à medida que realiza as tarefas do teste.

Também é sugerido que, por razões éticas e a fim de manter os testes em execução, você deve providenciar ajuda (adequada de acordo com o tipo de aplicação) se o participante estiver ficando desconfortável ou se ele travar completamente.

Quanta ajuda é necessária? Isso irá depender do tipo de aplicação. Por exemplo, no caso de um quiosque de informação para uso do público, você talvez precise de ajuda mínima. Também devemos considerar o grau de completude da aplicação de teste e, particularmente, se quaisquer recursos de ajuda foram implementados.

TEMA 5 – OUTRAS QUESTÕES SOBRE A AVALIAÇÃO

Finalizado, você deverá realizar uma sessão piloto e consertar quaisquer dificuldades imprevistas. Para auxiliar, é sugerido preparar um **plano** para orientar a avaliação. O plano especifica:

- Os objetivos da sessão de teste.
- Detalhes práticos.
- Número e tipo de participantes.
- Tarefas a serem realizadas com a definição de término bem-sucedido (essa seção também especifica que dados devem ser coletados e como eles serão analisados).

Uma questão relevante é relatar os resultados da avaliação de usabilidade para a equipe de *design*. Por mais competente e completa que seja a avaliação, ela só vale a pena se os resultados gerarem ações!

Mesmo que você seja, ao mesmo tempo, *designer* e avaliador, precisa de uma lista organizada de constatações, de modo a priorizar o trabalho de redesenho. Se você se reporta a uma equipe de *design* ou desenvolvimento, é crucial que ela possa ver imediatamente qual é o problema, a importância das suas consequências e, de preferência, o que precisa ser feito para resolvê-lo.

Com o advento da conectividade pela *Internet*, as pessoas podem participar de avaliações sem estarem fisicamente presentes. Se a aplicação é baseada na *Internet* ou pode ser instalada remotamente, instruções podem ser fornecidas para que os usuários executem as tarefas de teste e preencham e devolvam questionários na forma impressa ou não. Os questionários *on-line* e os métodos de *crowdsourcing* que descrevemos são adequados neste caso

5.1 RASTREAMENTO DO MOVIMENTO DOS OLHOS – *EYE TRACKING*

O rastreamento do movimento dos olhos pode mostrar a mudança de foco dos participantes em diferentes áreas da tela. Isso pode indicar quais as características de uma interface de usuário que atraíram a atenção e em que ordem, ou captar indicadores de padrões de olhar em maior escala.

O rastreamento dos olhos é muito popular entre os designers de *websites*, usado para destacar áreas da página mais olhadas e as ignoradas. O equipamento de rastreamento dos olhos é montado em capacete ou no monitor de computador. *Softwares* de rastreamento estão prontamente disponíveis para fornecer mapas da tela. Técnicas fisiológicas de avaliação partem do princípio de que todas as nossas emoções – ansiedade, prazer, apreensão, encanto, surpresa e assim por diante – geram mudanças fisiológicas.

As medidas mais comuns são mudanças no batimento cardíaco, no ritmo respiratório, na temperatura da pele, no volume de sangue bombeado e na resposta galvânica da pele (um indicador da quantidade de transpiração). Todos são indicadores de mudanças no nível geral de atenção, o que, por sua vez, indica uma reação emocional.

Sensores podem ser colocados no corpo do participante (normalmente nas pontas dos dedos) e ligados ao *software*, que converte os resultados para o formato numérico e gráfico com fins de

análise. Mas também existem muitos outros métodos discretos, como sensores de pressão na direção de um *videogame*, ou sensores que medem se o participante está sentado na beirada da cadeira.

Com relação à emoção evocada, não pode ser deduzida somente a partir do nível de atenção, como também a outros dados: expressões faciais, postura ou questionamento direto. Alguns podem também medir a dilatação da pupila, o que é considerado um sinal de interesse. A sua pupila se dilata se você gosta do que vê!

5.2 PRESENÇA

Outra aplicação corrente é a avaliação do grau de presença – a sensação de “estar lá” evocada pelos ambientes virtuais. Geralmente, eventos que assustam ou características ameaçadoras são produzidos no ambiente e os níveis de atenção medidos quando as pessoas se defrontam com eles

Pesquisadores conduziram uma série de experimentos e mediram a atenção à medida que os participantes se aproximam de um “precipício virtual”. Nessas circunstâncias, as alterações no batimento cardíaco estavam mais intimamente relacionadas aos autoinformes de estresse.

Na avaliação da presença, *Designers* de aplicações de realidade virtual – e de algumas aplicações multimídia também – estão regularmente preocupados com o senso de presença, de estar “lá” no ambiente virtual e não “aqui” na sala onde a tecnologia é usada.

Acredita-se que um forte senso de presença é crucial para aplicações como jogos, aplicativos destinados a tratar fobias, os que permitem às pessoas “visitarem” lugares de verdade ou mesmo para algumas aplicações de trabalho, tais como treinamento para operar com eficácia sob estresse. As dificuldades incluem:

- **O senso de presença** – Fortemente entremeado com disposições, experiências e expectativas individuais.
- **O conceito de presença** – Mal definido e é objeto de muito debate entre os pesquisadores.
- **A consciência da presença** – Perguntar às pessoas sobre presença enquanto estão passando pela experiência do ambiente virtual tende a interferir com a própria experiência. Por outro lado, fazer perguntas retrospectivamente inevitáveis deixa de capturar a experiência enquanto vivida.

As medidas usadas na avaliação de presença adotam várias estratégias para evitar esses problemas, mas nenhuma é totalmente satisfatória.

5.3 OUTRAS MEDIÇÕES

Tentativas menos estruturadas de captar relatos verbais de presença incluem pedir às pessoas que escrevam relatos da sua experiência, ou convidá-las a fornecer comentários livres em uma entrevista. Os resultados são, então, analisados quanto a indicações de um senso de presença.

Aqui, a dificuldade está em definir o que deve ser tratado como um indicador e, também, as **camadas de obliquidade** introduzidas pela relativa destreza verbal do participante e pela interpretação imposta pelo analista. Outras abordagens de medição de presença tentam evitar essas camadas de obliquidade, observando o comportamento no ambiente virtual ou fazendo medições fisiológicas diretas.

Em complemento, as pessoas em casa não são uma “audiência cativa” para o avaliador tanto quanto aquelas que estão no trabalho. É também mais provável elas estarem mais preocupadas em proteger sua privacidade e, geralmente, não estão dispostas a gastar seu tempo de lazer ajudando você com a sua avaliação de usabilidade.

Por isso, é importante que as técnicas de coleta de dados sejam interessantes e estimulantes para os usuários, demandando o mínimo possível de tempo e esforço. Este é um campo ainda em desenvolvimento e os pesquisadores continuam a adaptar abordagens existentes e a desenvolver novas abordagens.

Petersen, Madsen e Kjaer (2002) usaram entrevistas convencionais na época em que uma nova televisão havia sido instalada e fizeram o acompanhamento pedindo que as famílias interpretassem cenários de uso. Diários também foram distribuídos como ferramenta de coleta de dados. Mas, neste caso, a taxa dos que não o preencheram foi alta, possivelmente em função da complexidade do formulário e da incompatibilidade entre um diário particular e a atividade social de assistir televisão.

O *tour* tecnológico é um exemplo eficaz que é relatado. Aqui o investigador forneceu *post-its* aos usuários para captar seus pensamentos sobre os conceitos de *design* (recolhidos posteriormente). Uma ilustração de cada conceito diferente foi deixada na casa em um local onde poderia ser usada e os usuários foram estimulados a pensar em como eles usariam o dispositivo e em que tipos de

problemas poderiam surgir. As observações foram anotadas nos *post-its*, que foram, depois, grudados na ilustração e recolhidos mais tarde.

Quando a família é o foco de interesse, as técnicas devem ser interessantes para crianças e adultos – ajuda a garantir que todos os pontos de vista estejam cobertos e atrai os pais para as atividades de avaliação.

FINALIZANDO

Nesta aula, pudemos conhecer uma seleção diversificada de técnicas de avaliação que se adequam a diferentes circunstâncias. Algumas delas estão relativamente bem estabelecidas e outras ainda são objeto de pesquisas. No entanto, o que todas têm em comum é que elas que **baseiam a avaliação em verdadeiros contextos de uso**.

É provável que nenhuma técnica de avaliação se encaixe exatamente nas circunstância da sua própria avaliação; portanto, faça o que outros profissionais fazem: considere o que realmente tem de avaliar, revise o disponível e amplie ou adapte onde for necessário. Evidentemente, isso é mais eficaz quando se tem um bom conhecimento funcional, tanto dos aspectos práticos quanto dos aspectos teóricos do contexto da aplicação.

Diferentes objetivos de avaliação requerem que perguntas diferentes sejam respondidas. Tanto a revisão por especialista quanto o teste pelo usuário final são eficazes, mas ambos devem ser usados juntos, como métodos complementares.

Praticamente todos os graus de avaliações com participantes podem resultar em percepções úteis, mas deve-se tomar cuidado com generalizações quando o número de usuários for pequeno.

REFERÊNCIAS

BARBOSA, S. D. J. et al. **Interação Humano-Computador e Experiência do Usuário**. Leanpub, Victoria, Colúmbia Britânica, Canadá: 2021. Disponível em: <<http://leanpub.com/ihc-ux>>. Acesso em: 11 out. 2021.

BENYON, D. **Interação Humano Computador**. 2. ed. São Paulo: Pearson Prentice Hall, 2011.

CARVALHO, T. B. R.; **Usabilidade para Web Sites Institucionais**: estudo de caso do web site Estúdio 8bits. Disponível em: <<https://repositorio.ufpe.br/bitstream/123456789/31636/1/CARVALHO%2C%20Thiago%20Bezerra%20Regis%20de.pdf>>. Acesso em: 11 out. 2021.

PREECE, J.; ROGERS, Y.; SHARP, H. **Design de Interação**. Bookman Editora, 3. ed., 2013.

PREECE, A. D. et al. **Human-machine conversations to support mission-oriented information provision**. 2013. Disponível em: <<https://doi.org/10.1145/2509338.2509342>>. Acesso em: 11 out. 2021.

SHEDROFF, N. **Experience Design** 1. ed. Indianapolis: New Riders, 2001.

TIGER, L. **The Pursuit of Pleasure**. 1. ed. Transaction Publishers, 2000.