

**"Software is not usable.**

**Instead, *software gets used*, and the resulting user experiences are a composite of several qualities that are shaped by product attributes, user attributes and the wider context of use."**

Gilbert Cockton

"Usability Evaluation"

Chapter 15 - Encyclopedia of HCI



**IME-USP**

# **MAC0446/MAC5786**

## **Introdução à IHC**

### Avaliação

Prof. Carlos Hitoshi Morimoto  
Departamento de Ciência da Computação - IME/USP  
<http://www.ime.usp.br/~hitoshi>

# Leituras

- Encyclopedia of Human-Computer Interaction
  - Chapter 15: Usability Evaluation  
By Gilbert Cockton

## Recomendados:

- Capítulo 13 do livro Preece & Rogers  
“Introducing Evaluation”

# The language of evaluation

Analytics

Analytical evaluation

Controlled experiment

Expert review or crit

Field study

Formative evaluation

Heuristic evaluation

In the wild evaluation

Living laboratory

Predictive evaluation

Summative evaluation

Usability laboratory

User studies

Usability testing

Users or participants

# Objetivos dessa aula

- Discutir como os desenvolvedores lidam com restrições do mundo real
- Explicar os conceitos e termos utilizados para discutir avaliação de interfaces
- Examinar como técnicas distintas são utilizadas em estágios diferentes do desenvolvimento
- Apresentar algumas técnicas usadas para avaliação de interfaces
- Discutir temas conceituais, práticos e éticos que devem ser considerados ao planejar uma avaliação

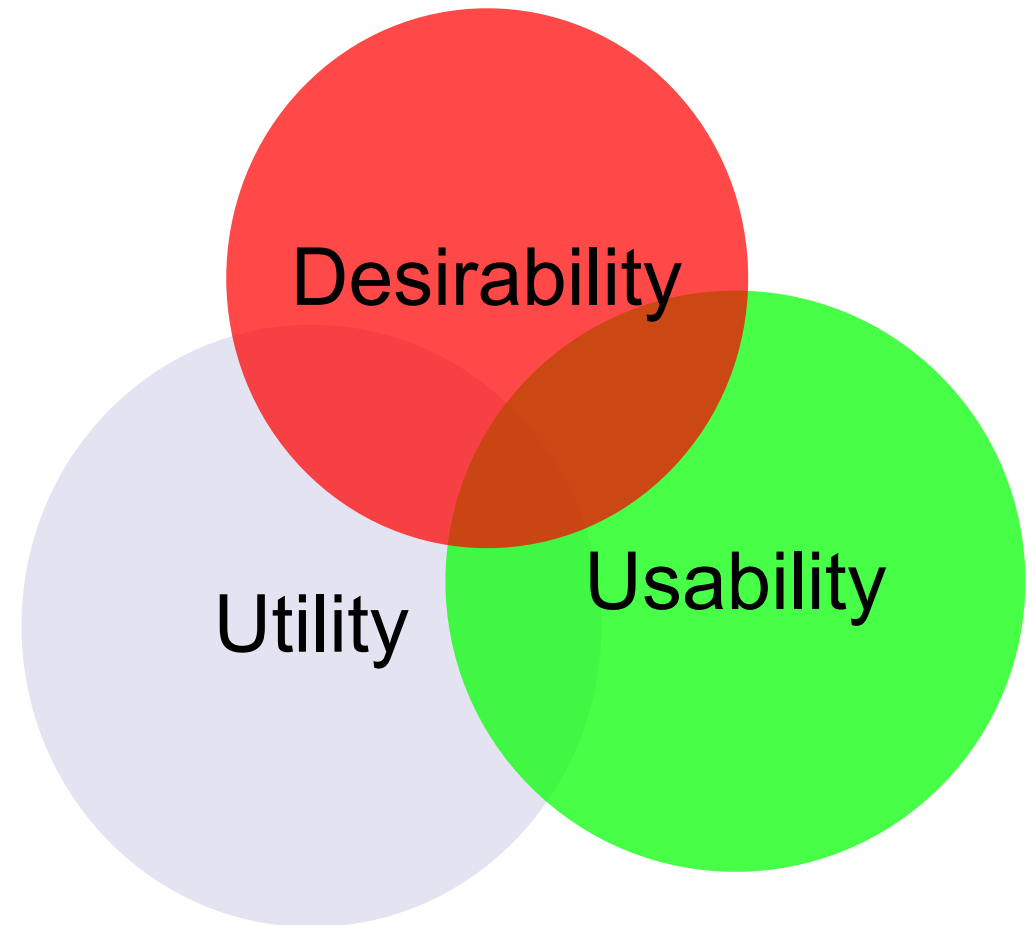
# Tópicos

- O que é avaliar?
  - o que a gente deve avaliar no design?
- Por que avaliar?
- Quando avaliar?
- Como avaliar?

**O QUE AVALIAR?**

# Utilidade, Usabilidade e Desejabilidade

- Utilidade (o que)
  - Funcionalidades
  - Produto
- Usabilidade (como)
  - Efetivo, eficaz, seguro, etc.
  - Interaction design
- Desejabilidade (porque)
  - Satisfação
  - UX design
- Essa classificação nos ajuda a recordar todos os objetivos que devem ser almeçados no design





# Objetivos de usabilidade

- Eficácia, eficiência, segurança, utilidade, facilidade de aprender e lembrar.
- Uma tecnologia pode ser avaliada aplicando esses objetivos.
  - O quanto ela satisfaz esses objetivos?
- Requer cenários específicos de uso para serem efetivos.
  - ajudam a definir o contexto de uso

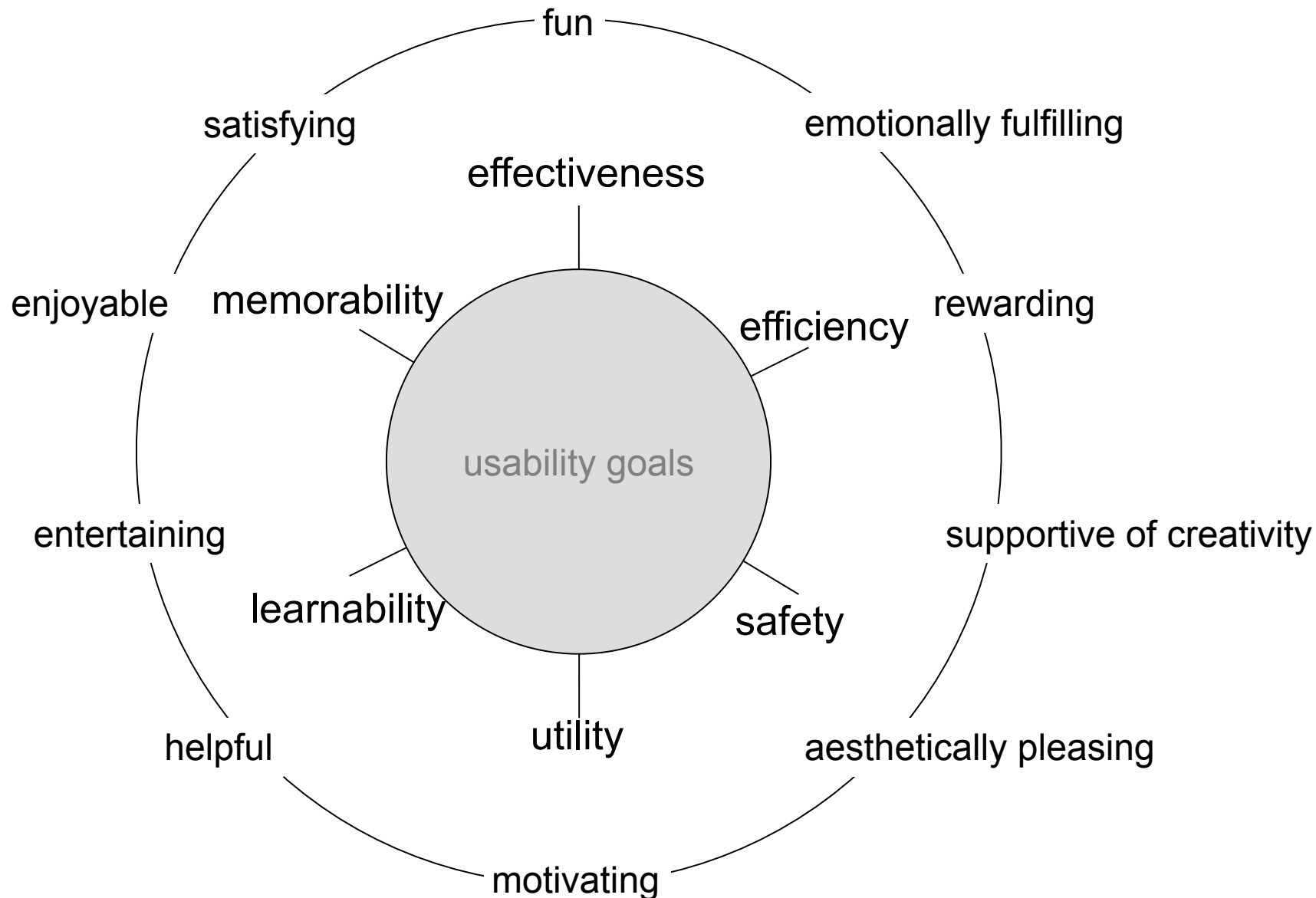
# Princípios de usabilidade

- Don Norman: visibilidade, feedback (retorno), restrições (físicas, lógicas, culturais), mapeamento, consistência, affordance.
  - Também conhecidos como princípios de design
- Dix e outros:
  - facilidade para aprender: prever o comportamento, sintetizar, familiaridade, generalizar, consistente.
  - Flexibilidade: diálogo intuitivo, multi-threading, migração de tarefas, substitutividade, customizável
  - Robustez: observável, recuperação de erros, tempo de resposta, conformidade com a tarefa.
- Novamente, aplicações precisam de cenários específicos de uso para avaliação ser efetiva.

# Experiência do Usuário



# Como os objetivos de usabilidade se relacionam com os objetivos de experiência?



**POR QUE AVALIAR?**

# Por que avaliar?

- Comparar o sistema com concorrentes
- Identificar oportunidades para inovação: soluções p/ problemas que os usuários tem com os sistemas atuais ou novos
- Descobrir se a produtividade e usabilidade melhoram
- Determinar se o sistema atinge os níveis de desempenho/segurança determinados por lei

# Por que avaliar tecnologias existentes?

- Para desenvolver um conjunto de **técnicas analíticas** que ajudem você a definir se a tecnologia é boa ou não
- Para desenvolver suas **habilidades críticas** que serão úteis nas suas próprias atividades de design
- Para permitir que você seja capaz de **sugerir mudanças** no design baseadas em análises estruturadas

# Por que avaliar?

- o design para interação, através de seu ciclo de design e teste iterativo, é a única metodologia conhecida já validada, capaz de produzir bons resultados de forma consistente. Se você não tiver testes com usuários como uma parte integral de seu processo de design, você vai desperdiçar muito dinheiro.

Bruce Tognazzini

Veja [www.asktog.com](http://www.asktog.com) para outras discussões a respeito de design e avaliação de interfaces





# Os argumentos do Tog

- Os problemas são resolvidos ANTES do produto ser lançado, e não depois
- O time pode se concentrar nos problemas verdadeiros, e não imaginários
- Os engenheiros podem codificar ao invés de debater
- Tempo para lançamento é reduzido
- Um design robusto já no lançamento

**QUANDO AVALIAR?**

# Quando avaliar

- Projeto e avaliação iterativa é um processo contínuo que examina:
  - ideias para o modelo conceitual
  - protótipos do novo sistema
  - com o tempo, protótipos mais maduros e completos

Projetistas precisam verificar se eles entenderam os requisitos dos usuários

# Quando avaliar?

- Ao longo de todo o processo de design
- A partir das primeiras descrições, sketches, etc., das necessidades dos usuários até o produto final
- O processo de design continua através de ciclos iterativos de 'design - teste - redesign'
- Avaliação é o principal ingrediente para um design de sucesso

**COMO AVALIAR?**

# Estudo de usuários (user studies)

- Estudo de usuários envolve observar o comportamento das pessoas em seu ambiente natural, ou em laboratório, tanto com tecnologias antigas quanto com modernas.
  - Abigail Sellen

# Tipos de Avaliação

- Ambientes controlados envolvendo os usuários: testes de usabilidade, experimentos em labs e living labs
- Ambientes naturais com usuários: estudos de campo para ver como o produto é usado no mundo real
- Qualquer ambiente não envolvendo usuários: críticas de consultores; heurísticas; para prever, analisar e modelar aspectos da interface

# laboratório de usabilidade

- Descrição:
  - gravação da atividade de usuários típicos realizando tarefas típicas.
  - Em ambientes controlados, pode ser realizada em laboratórios ou em campo.
  - A gravação pode ser em vídeo, ou seqüência de teclas, tempo para completar tarefas, etc.





# Teste de usabilidade

- Vantagens
  - sem interrupções
  - pode avaliar desempenho, identificar erros e ajudar a explicar porque os usuários fizeram o que fizeram
  - questionários sobre satisfação e entrevistas
- Desvantagens
  - falta de contexto, difícil determinar usuário/tarefa típicos
  - Tempo para preparar os testes, recrutar participantes, e realizar os testes
  - necessita de recursos/equipamentos

# Estudos de campo

- **Descrição:**
  - observações e entrevistas nos ambientes naturais
- **Vantagens:**
  - ajuda a entender o que os usuários fazem naturalmente e qual o impacto da tecnologia em seu contexto
  - para o design: identifica oportunidades, determina requisitos, decide a melhor forma de introduzir a tecnologia, avaliação em uso real
- **Desvantagens**
  - acesso aos locais
  - falta de controle, ruído, distrações

# Living Labs

- O uso de tecnologia no dia-a-dia das pessoas pode ser avaliado em "living labs"
- Essas avaliações são difíceis de se fazer em usability labs
- Exemplo: a "Aware Home" foi populada com uma rede de sensores e câmeras (Abowd et al., 2000) .

# Avaliação sem usuários

- Avaliação heurística
- Caminhada cognitiva
- Data analytics

# Avaliação Ligeira (Quick & Dirty)

- Avaliação ligeira descreve a forma comum através da qual os designers obtém feedback informal dos usuários para confirmar que suas idéias estão de acordo com as necessidades dos usuários
- Avaliações ligeiras são feitas a qualquer momento
- O objetivo seria um teste rápido para o processo de design, em contraste com um estudo com conclusões bem documentadas
- Desvantagens: não necessariamente precisas e sem documentação cuidadosa sobre descobertas

# Avaliação preditiva

- Descrição
  - perguntar a experts para aplicar o conhecimento de usuários típicos
    - uso de personas e cenários
  - aplicar modelos teóricos
    - GOMS (goals, operators, methods, selection rules)
- Vantagens
  - ajuda a prever problemas de usabilidade e como usuários irão realizar certas tarefas
- Desvantagens
  - não há envolvimento direto com usuários, ou um contexto real

# Técnicas de avaliação

O envolvimento dos usuários varia conforme a técnica.

- Observar os usuários (lab ou campo)
- Perguntar a opinião de usuários
- Perguntar a opinião de experts
- Testar (medir) o desempenho de usuários
- Modelar o desempenho das tarefas de usuários

# Técnicas de avaliação

- **Avaliações objetivas:** mensuram ou modelam o desempenho de tarefas típicas
  - fluxo (throughput): medida de produtividade (ex: número de comandos usados, menus navegados, janelas abertas, etc)
  - tempo de execução: tempo necessário para realizar as operações. Ex: tempo para selecionar uma opção, tempo para recuperar de um erro
  - precisão: ex: seleção de opção menu tipo pizza x lista
  - erro: número, quando, onde, severidade, frequência
- **Avaliações subjetivas:** mensuram opiniões de usuários
  - opiniões e preferências medidas através de votações, questionários e entrevistas
  - observações e impressões dos avaliadores



# Exemplos de técnicas

- Caminhada cognitiva (expert)
- Avaliação heurística (expert)
- caminhadas de procedimentos (expert)
- cenários baseados nas tarefas (user test)
- pensar em voz alta (ligeiro, user test)
- observação (campo, ligeiro, user)
- caminhada após a tarefa (campo, user)
- entrevistas (campo, ligeiro, user)
- grupos de foco
- questionários
- rastreamento de olhar, etc

# Qual técnica escolher?

- A seleção da técnica depende dos objetivos de um determinado estágio do processo de design
  - métodos podem variar de formais (ex: benchmark) a informais (grupos de foco). Ex:
    - critica de um sistema existente (ex: website)
      - avaliação heurística
    - satisfaz objetivos específicos de usabilidade
      - benchmark
    - escolha entre designs alternativos
      - experimento controlado em labs
    - busca de informação nos primeiros estágios de design
      - cenários, estudo de campo

# Métodos de avaliação

Método	Ambiente controlado	Ambiente natural	Sem usuários
Observação	X	X	
Perguntas aos usuários	X	X	
Perguntas aos experts		X	X
Testes	X		
Modelamento			X

# O arcabouço DECIDE

- Determinar os objetivos da avaliação
- Explorar questões específicas que precisam ser respondidas
- (Choose) Escolha o paradigma e técnica para responder as questões
- Identificar as considerações práticas
- Decidir como lidar com as considerações éticas
- (Evaluate) Avaliar, interpretar e apresentar os dados

# Determinar os objetivos

- Quais os objetivos de alto nível?
- Quem o deseja e por que?
- Os objetivos ajudam a determinar o paradigma
- Exemplos:
  - identificar a melhor metáfora para utilizar no design
  - verificar se a interface final é consistente
  - investigar como a tecnologia afeta o ambiente de trabalho
  - melhorar a usabilidade de um produto existente

# Explorar as questões

- As avaliações precisam de objetivos e questões para guiá-las e assim não desperdiçar tempo
- exemplo: descobrir porque usuários preferem comprar bilhetes em papel a e-tickets.
  - qual a atitude dos clientes com relação a e-tickets
  - eles se preocupam sobre segurança?
  - a interface para a sua obtenção é pobre?
- Exercício: que perguntas você faria sobre o projeto de um telefone celular?

# Escolha de paradigma e técnica

- O paradigma de avaliação influencia a técnica a ser utilizada, e como os dados são analisados e apresentados
- exemplo: estudo de campo não envolve teste e modelamento

# Identificar considerações práticas

- exemplo: como...
  - selecionar usuários
  - se manter dentro do orçamento
  - se manter no cronograma
  - encontrar avaliadores
  - selecionar equipamentos
  - reservar local



# Considerações éticas

- Desenvolver um formulário de consentimento
- participantes tem o direito de:
  - conhecer os objetivos do estudo
  - o que vai ser feito das descobertas
  - privacidade sobre as informações pessoais
  - não ser considerado sem consentimento
  - podem sair quando desejar (mesmo após coleta)
  - serem tratados polidamente

# Avaliar, interpretar e apresentar os dados

- Como os dados são analisados e apresentados depende do paradigma e das técnicas utilizadas
- isso também deve ser considerado:
  - consistente ou confiável: os resultados podem ser replicados?
  - validade: está se medindo o que foi pensado?
  - vícios: o experimento é viciado?
  - escopo: os resultados podem ser generalizados?
  - validade ecológica: o ambiente está afetando o estudo?

# Estudos piloto

- um pequeno experimento do estudo principal
- certificar que o plano é viável
- verificar:
  - que o procedimento pode ser realizado
  - que os scripts das entrevistas, questionários, experimentos, etc, funcionam
- vale a pena fazer vários testes piloto para eliminar problemas antes do estudo principal
- peça aos seus colegas quando você não puder utilizar usuários

# Pontos importantes

- a avaliação e design seguem juntos
- não avaliar pode significar um grande prejuízo
- um paradigma de avaliação é uma metodologia que é influenciada por teorias e filosofias particulares
- 5 categorias de técnicas foram identificadas: observação de usuários, questionando usuários, questionando experts, testando usuários, e modelando usuários
- o arcabouço DECIDE: determinar objetivos, explorar as questões para atingir objetivos, escolher o paradigma/técnica, identificar questões práticas e éticas, avaliar o resultado e apresentar o resultado
- realize estudos piloto antes do experimento principal