

2 Fundamentação Metodológica

Este capítulo descreve os fundamentos metodológicos utilizados neste estudo, bem como trabalhos de avaliação de IHC relacionados a este estudo, mais especificamente os que utilizam o método de avaliação de comunicabilidade (MAC) e métodos de avaliação de usabilidade aplicados a ambientes virtuais.

2.1. Avaliação de IHC

Para que se obtenha melhores resultados, a avaliação de IHC deve ser realizada ao longo de todas as fases de desenvolvimento de um sistema de computador.

Como mencionado anteriormente, este estudo trata da avaliação da HybridDesk [Carvalho 2009], cuja avaliação de IHC traz uma questão desafiadora de qual método, ou métodos, deveriam ser utilizados.

Se considerarmos que as funcionalidades de um sistema são implementadas pelos desenvolvedores, que são em geral os próprios designers do sistema, a compreensão da intenção do designer torna-se de suma importância para uma avaliação correta da interação do usuário com o sistema. Essa mensagem de metacomunicação emitida pelo designer é exatamente o foco principal da avaliação de comunicabilidade.

Portanto, usabilidade e comunicabilidade são fatores distintos que contribuem para a qualidade da experiência do usuário. Em função disso, a avaliação da HybridDesk foi realizada aplicando-se o método de avaliação de comunicabilidade (MAC) e diversos métodos de avaliação de usabilidade, que no nosso caso foram: avaliação heurística, observação de uso com *talk-aloud*, e questionário e entrevista pós-uso.

O principal foco das avaliações foi a identificação de questões de comunicabilidade e usabilidade que afetam a interação na HybridDesk. Tullis & Albert (2008) definem questões de usabilidade como:

- Qualquer coisa que impede a finalização da tarefa;
- Qualquer coisa que leva o usuário a um caminho errado;

- Qualquer coisa que cria algum nível de confusão;
- Qualquer coisa que produz um erro;
- Não ver alguma coisa que deveria ser vista;
- Assumir que alguma coisa está correta quando não está;
- Assumir que uma tarefa está completa quando não está;
- Executar a ação errada;
- Interpretar errado alguma parte do conteúdo;
- Não entender a navegação.

2.2.

O Método de Avaliação de Comunicabilidade (MAC)

O método de avaliação de comunicabilidade (MAC) é baseado na teoria de engenharia semiótica [de Souza 2005], trazendo para o mesmo contexto comunicativo os designers, usuários e sistemas. Esta teoria caracteriza IHC como um tipo específico de metacomunicação mediada por computador [de Souza 2005, Salgado 2007].

O MAC foi proposto por Prates e coautoras (2000a) e utilizado para avaliação de editores HTML pelas autoras (2000b). Foi também documentado em livro por de Souza (2005), na dissertação de Salgado (2007) e em livro de de Souza & Leitão (2009), além de citado por Sharp e coautoras (2007).

O método pode ser aplicado em diferentes fases do design. Neste estudo aplicaremos o MAC na avaliação de um protótipo em fase final de desenvolvimento.

Salgado (2007) explica em sua dissertação que “o MAC é realizado através de cinco passos: preparação do teste, aplicação do teste, etiquetagem, interpretação e elaboração do perfil semiótico.”

As seções a seguir apresentam uma descrição sumarizada das etapas para a realização do MAC, com base na descrição detalhada apresentada na dissertação de Salgado (2007).

2.2.1 Preparação do Teste

Esta etapa tem como objetivo definir todos os parâmetros gerais do teste e preparar todo o conteúdo do teste. Ela envolve as seguintes atividades [Salgado 2007, pp. 30-36]:

- a. Estudo da documentação disponível para a aplicação. Uma entrevista com o designer pode suprir a falta de documentação;
- b. Realização de uma inspeção informal do design do software, especialmente dos signos estáticos e dinâmicos com enfoque nos aspectos da comunicação, anotando os problemas potenciais. Se disponível, confirmar as intenções de design com o próprio designer;
- c. Elaboração de uma versão consolidada da metacomunicação e das questões de comunicação identificadas individualmente pelos avaliadores;
- d. “Definição do perfil dos participantes do teste. O perfil deverá ser compatível com a metacomunicação designer-usuário identificada pelos avaliadores na atividade de inspeção. De Souza (2005) recomenda que os participantes (usualmente de 3 a 10) representem o ‘usuário típico’ da aplicação”.
- e. Elaboração de cenários de uso relevantes para a fase de observação de uso. A seleção dos cenários e atividades devem ser compatíveis com as questões de design identificadas na inspeção informal;
- f. Definição dos avaliadores de acordo com as recomendações sobre o perfil necessário;
- g. Análise das questões éticas envolvidas no teste;
- h. Elaboração dos roteiros das entrevistas e/ou questionários pré-testes;
- i. Elaboração do roteiro das entrevistas pós-uso;
- j. Elaboração do roteiro da observação de uso;
- k. Preparação do Termo de Consentimento para os participantes;
- l. Preparação do material para a observação de uso;
- m. Configuração do equipamento e da aplicação;
- n. Realização de um teste piloto.

2.2.2 Aplicação do Teste

Esta etapa corresponde à observação de uso em outros métodos de avaliação de IHC. “Esta etapa tem como objetivo a captura de evidências de uso da aplicação para posterior análise na fase de etiquetagem” [Salgado 2007].

É recomendado que esta fase tenha a participação de dois avaliadores que tenham conhecimento do MAC [Salgado 2007].

Esta fase envolve as seguintes atividades [Salgado 2007, pp.36-40]:

- a. Recepção do participante;
- b. Encaminhamento do participante ao sistema onde ocorrerá a observação de uso;
- c. Preparação dos equipamentos para a observação de uso;
- d. Apresentação do cenário de teste e início da gravação;
- e. Anotações dos avaliadores durante a observação de uso;
- f. Finalização da observação de uso;
- g. Finalização da fase de aplicação da avaliação.

2.2.3 Etiquetagem da Interação

A etiquetagem é a primeira etapa analítica do método, onde os avaliadores utilizam os produtos das etapas anteriores para a geração da interpretação [Salgado 2007].

Na etapa de etiquetagem o avaliador examina as quebras de comunicação designer-usuário e, a partir daí, infere certos aspectos de comunicabilidade. “O comportamento do usuário durante a interação é analisado e expressões de comunicabilidade, também chamadas de etiquetas, são usadas para representar a reação do usuário ao que acontece durante a interação” [Salgado 2007].

A etiquetagem das quebras de comunicação pode ser entendida como se o avaliador estivesse “colocando palavras na boca do usuário” por meio das expressões de comunicabilidade [de Souza 2005, Prates et al. 2000b].

O MAC utiliza atualmente 13 expressões básicas de comunicabilidade, apresentadas na Tabela 1 [Salgado 2007, pp.40-51].

Etiqueta	Descrição
“Cadê?”	“Esta etiqueta é usada quando o usuário conhece a operação que deseja executar, mas não a encontra de imediato na interface. Isto é equivalente a saber o que dizer, mas não encontrar palavras para dizê-lo.”
“Para mim está bom.”	“Esta etiqueta é tipicamente usada quando o usuário está convencido de que alcançou o seu objetivo. Esta etiqueta também pode ocorrer como uma espécie de “modalizador” (problemático, por certo) da forma de atingir o objetivo, sinalizando que o usuário entendeu errado (ou não entendeu) uma mensagem do designer ao longo da interação, e o problema persistiu até o final da sessão.”
“Assim não dá.”	“Esta etiqueta é usada onde o usuário abandona o caminho de interação composto de vários passos, porque ele pensa que esta opção não o está levando para seu objetivo. Isto é equivalente a constatar que o interlocutor não está entendendo nada do que está sendo dito, e assim tomar uma outra linha de comunicação. O sintoma típico é quando o usuário de repente interrompe uma atividade e segue uma direção totalmente diferente.”

“Ué, o que houve?”	“Esta etiqueta é usada quando o usuário repete uma operação porque não pôde ver o resultado dela. O sintoma típico é o usuário repetir a ativação de uma função a qual o feedback não existe ou não foi percebido pelo usuário.”
“Por que não funciona?”	“Esta etiqueta é usada quando o usuário não entende ou não se conforma com o fato de a operação efetuada não produzir o resultado esperado. O sintoma típico é o usuário repetir a ação. O uso desta expressão normalmente acontece porque o usuário acredita que o que está fazendo deveria causar os efeitos desejados. Por isso o usuário insiste na mesma ação tentando descobrir se uma pequena mudança de parâmetros ou contexto é possível ou necessária para fazer a ação funcionar.”
“O que é isto?”	“Esta etiqueta é usada quando o usuário espera ver alguma dica explicativa ou algum outro tipo de indicação sobre o significado de um signo da interface. O sintoma típico é quando o usuário posiciona o cursor do mouse em algum signo da interface, esperando que uma dica seja apresentada. Também envolve a exploração de menus, listas de <i>pull-down</i> e caixas de diálogos para ver o que “dizem”.”
“Socorro!”	“Esta etiqueta é usada quando o usuário não consegue realizar sua tarefa através da exploração da interface. O sintoma típico é quando o usuário ativa a função de help, procura por documentação on-line ou off-line ou pede explicações a alguma pessoa.”
“Epa!”	“Esta etiqueta é usada quando o usuário executa uma ação indesejada e, percebendo imediatamente que isto ocorreu, desfaz a ação. O sintoma típico é quando o usuário imediatamente aciona o <i>undo</i> ou cancela um quadro de diálogo aberto indevidamente.”
“Onde estou?”	“Esta etiqueta é usada onde o usuário está interpretando (e potencialmente usando) signos em um contexto errado da aplicação. É como se o usuário estivesse dizendo uma coisa que faz parte do sistema de significação da interface, mas está dizendo na hora ou no local errado. O sintoma típico é quando o usuário tenta executar operações ou busca por signos em um modo, enquanto está em outro.”
“E agora?”	“Esta etiqueta é usada quando o usuário não sabe o que fazer e procura descobrir qual o seu próximo passo. Por isso ele sequer consegue formular uma intenção de comunicação. O sintoma típico é vagar com o cursor sobre a interface e inspecionar os menus de forma aleatória ou sequencial.”
“Vai de outro jeito.”	“Esta etiqueta é usada quando o usuário não consegue realizar a tarefa da forma como o projetista idealizou, e resolve seguir outro caminho, geralmente mais longo ou complicado. O sintoma típico é quando o usuário atinge o seu objetivo de uma maneira não ótima.”
“Não, obrigado.”	“Esta etiqueta é usada onde o usuário conhece a solução preferencial do designer, mas opta explicitamente por uma outra forma de interação. O sintoma típico consiste da ocorrência da ação preferencial disponibilizada pelo designer, seguida de uma ou mais formas alternativas de se alcançar o mesmo resultado.”
“Desisto.”	“Esta etiqueta é usada onde o usuário explicitamente admite sua incapacidade em alcançar seu objetivo. Ela pode ocorrer durante a interação, quando o usuário desiste de alguma atividade, mas continua executando os próximos passos. Assim como pode ocorrer no final da interação quando o usuário desiste da tarefa como um todo. O sintoma típico é a desistência de tarefas intermediárias ou a interrupção prematura da tarefa.”

Tabela 1 – Etiquetas Usadas na Avaliação de Comunicabilidade [Salgado 2007, pp.40-

2.2.4 Interpretação dos Dados

Esta etapa deve ser realizada por um avaliador, ou avaliadores, com especialização em IHC, de preferência com bons conhecimentos de engenharia semiótica, uma vez que os resultados obtidos são dependentes das habilidades e conhecimentos dos avaliadores [Salgado 2007].

Salgado (2007) explica que “a análise das ‘falas’ de comunicabilidade permite ao avaliador algumas indicações das causas e provavelmente possíveis soluções para o problema identificado.”

O avaliador deve analisar e organizar (ou classificar) as evidências coletadas pela fase de etiquetagem de acordo com quatro diferentes perspectivas, que gradualmente o levam a níveis mais abstratos de interpretação dos dados empíricos [de Souza 2005, de Souza & Leitão 2009]:

- a. A frequência e contexto da ocorrência de cada tipo de etiqueta;
- b. A existência de padrões de sequências de tipos de etiquetas;
- c. O nível de problemas indicados pela ocorrência de tipos de etiquetas e sequência; e
- d. As questões de comunicabilidade que causaram as rupturas de comunicação observadas.

A interpretação pode ser ainda enriquecida através de mapeamentos específicos, como por exemplo, uma classificação de problemas de interação: execução, navegação, atribuição de significado e percepção (vide Tabela 2) [Salgado 2007].

Expressões de Comunicabilidade (Etiqueta)	Execução	Navegação	Atribuição de Significado	Percepção
“Cadê?”		X		
“Para mim está bom.”			X	X
“Assim não dá.”		X	X	X
“Ué, o que houve?”			X	X
“Por que não funciona?”			X	X
“O que é isto?”			X	
“Socorro!”	X	X	X	X
“Epa!”		X	X	
“Onde estou?”		X	X	X
“E agora?”		X	X	X
“Vai de outro jeito.”	Incompreensão/desconhecimento da solução preferencial do designer			
“Não, obrigado.”	Recusa de usar a solução preferencial do designer			
“Desisto.”	X			

Tabela 2 – Alinhamento das Categorias Conversacionais com Problemas Interativos de Alto Nível e de Usabilidade [de Souza et al. 1999, Salgado 2007]

2.2.5 Elaboração do Perfil Semiótico

Esta é a última etapa do MAC, com a geração de um relatório que procura recuperar o significado da metacomunicação designer-usuário.

O perfil semiótico "ajuda o avaliador a identificar, explicar e informar as sugestões de redesign das interações problemáticas" [Salgado 2007, pp. 55-56]. Ele vai além das rupturas de comunicabilidade e problemas de interação identificados e cuida de um nível mais abstrato, a linguagem da interface [Prates et al. 2000a, Salgado 2007].

O especialista em engenharia semiótica revê os resultados e problemas identificados na fase anterior e emite um diagnóstico sobre a qualidade da recepção da metacomunicação. Esta importante atividade de revisão guiada pelos resultados das fases anteriores é o que diferencia o MAC de um teste de

usabilidade, por exemplo. Nesta revisão, o avaliador deve ir além das ocorrências identificadas no teste, e investigar possibilidades de problemas de comunicabilidade relacionados com essas ocorrências. [Salgado 2007].

A mensagem do designer poderá ser então explicitada pelo avaliador, que será capaz de assumir a primeira pessoa e responder às seguintes questões [de Souza 2005], conforme traduzido por Salgado (2007):

- a. “No meu entendimento, quem são (ou serão) os usuários do produto do meu design?”
- b. “O que eu aprendi sobre as necessidades e desejos destes usuários?”
- c. “No meu entendimento, quais são as preferências destes usuários com respeito a seus desejos e necessidades, e por quê?”
- d. “Portanto, qual sistema eu desenhei para estes usuários, e como eles podem ou devem usá-lo?”
- e. “Qual é a minha visão de design?”

2.3. Métodos de Avaliação de Usabilidade

Pesquisou-se diversas fontes de referência para identificar os métodos de avaliação mais adequados para aplicação neste estudo, levando em consideração o sistema a ser avaliado, a HybridDesk, e seu estágio de seu desenvolvimento, um protótipo em estágio avançado. Vale notar que alternativas de design para a interação foram avaliadas no desenvolvimento inicial da HybridDesk, no contexto de uma Tese de Doutorado [Carvalho 2009].

Portanto, este estudo assumiu que o designer já havia concluído o design do sistema, levando em conta outros atributos para a aceitação do sistema, como é o caso da utilidade do sistema [Nielsen 1993] para a tarefa pretendida. Por isso, além do MAC, o estudo focou a realização da avaliação de usabilidade através da aplicação de uma combinação de métodos, visando uma avaliação mais abrangente da interação com a HybridDesk, que produzisse resultados para o ajuste do design.

A Engenharia de Usabilidade foi proposta por Nielsen (1993) em seu livro que se tornou referência para diversos estudos e publicações. Segundo ele, a usabilidade tem múltiplos componentes e está tradicionalmente associada com cinco atributos:

- Facilidade de aprendizado: o sistema deve ser fácil de aprender para que o usuário possa rapidamente começar a produzir;

- Eficiência: o sistema deve ser eficiente de se usar, tal que uma vez o usuário tenha aprendido a usar o sistema, consiga atingir alta produtividade;
- Facilidade de memorizar: o sistema deve ser fácil de se lembrar, tal que o usuário casual possa retornar ao sistema um tempo depois sem ter de aprender tudo novamente;
- Poucos Erros: o sistema deve ter uma baixa taxa de erros, tal que os usuários cometam poucos erros ao usar o sistema;
- Satisfação: o sistema deve ser agradável ao uso, tal que os usuários fiquem subjetivamente satisfeitos ao usar o sistema.

Ele também identifica diversos métodos de avaliação de usabilidade:

- Avaliação Heurística;
- Medidas de Desempenho;
- *Thinking aloud*;
- Observação;
- Questionários;
- Entrevistas;
- *Focus Groups*;
- *Log* da utilização;
- *Feedback* do usuário.

Com o argumento de que “o melhor é o inimigo do bom”, Nielsen (1993) propõe uma metodologia que chamou de *Discount Usability Engineering*, visando reduzir ao máximo o custo da aplicação da avaliação de usabilidade. Esta metodologia se baseia nos seguintes métodos e técnicas:

- Observação do usuário e tarefas;
- Cenários;
- *Thinking-aloud* simplificado;
- Avaliação Heurística.

A “Observação do usuário e tarefas” envolve a visita aos usuários em seu ambiente de trabalho e a observação de como fazem uso do sistema, sem interferências. Sharp e coautoras (2007) denominam este método de “Estudos de campo”.

“Cenários” são usados para reduzir o nível de funcionalidade e o número de características do sistema sendo avaliadas. Eles podem ser implementados tanto como *mock-ups* de papel quanto como ambientes de prototipação.

De acordo com Nielsen (1993), o método de *thinking aloud* “envolve termos pelo menos um usuário de cada vez utilizando o sistema, para execução

de um dado conjunto de tarefas, ao mesmo tempo em que é solicitado a pensar alto”.

Nielsen (1993) propõe a utilização de um método simples de *think aloud*, onde a análise dos dados é feita com base em anotações feitas pelo avaliador, dispensando a necessidade de gravação da sessão em vídeo, como requerida pelo método tradicional de *think-aloud*.

Em seu livro, Nielsen (1993) escreve ainda que “a avaliação heurística é feita olhando-se para uma interface e tentando-se emitir uma opinião sobre o que é bom ou ruim nesta interface”.

O método de *think-aloud* foi proposto por Ericsson & Simon (1993), que fornecem uma explicação detalhada sobre o uso de “relatos verbais como dados”, produzindo a base para a aplicação do método de usabilidade *think-aloud* durante uma sessão de observação de uso.

Vale notar que eles identificaram e definiram três níveis diferentes de verbalização:

- Nível 1: “é simplesmente a vocalização de codificações verbais, como requerido pelas tarefas. (...) Nesse nível, não existem processos intermediários, e o usuário não precisa fazer qualquer esforço especial para comunicar seus pensamentos.” (p.79)
- Nível 2: “envolve a descrição, ou melhor, a explicação do conteúdo do pensamento. (...) [Inclui] “verbalizações que não trazem novas informações para o foco de atenção do usuário.” (p.79) Uma vez que exige um certo tempo de processamento, é esperado que vá introduzir algum atraso, mas não muda a estrutura do processo para executar a tarefa principal.
- Nível 3: “requer que o usuário explique seus processos de pensamento ou pensamentos.” (p.79) Isto não apenas adiciona um atraso, como pode também alterar a estrutura para execução da tarefa principal.

Sharp e coautoras (2007) dedicaram alguns capítulos de seu livro à avaliação de sistemas interativos, incluindo a descrição de “Teste de Usabilidade” e “Estudos de Campo”. “Teste de Usabilidade” trata da observação do usuário em ambiente controlado, na maior parte das vezes num laboratório para realização de avaliações de desempenho. Nielsen (1993) chama este método de “Medidas de Desempenho”. Já “Estudos de Campo” trata da observação do usuário em seu ambiente de trabalho, como é o caso do método de “Observação do Usuário e Tarefas” descrito por Nielsen (1993).

Sharp e coautoras (2007) também tratam de métodos e técnicas de coleta de dados, incluindo:

- Gravação de dados;
- Entrevistas;
- *Focus groups*;
- Questionários;
- Observação (no campo, no laboratório e indireta).

Kuniavsky (2003) trata de aspectos da avaliação qualitativa da experiência do usuário, incluindo aspectos de usabilidade.

Para a realização de uma avaliação rápida e de baixo custo, ele propõe o que chama de teste de usabilidade "com amigos e família", o qual seria uma versão ainda mais simplificada do método *discount usability engineering* proposto por Nielsen (1993), pois não inclui uma avaliação heurística nem o uso do *think-aloud* durante a observação.

O método envolve os seguintes passos:

- Defina a audiência e seus objetivos;
- Crie tarefas que endereçam esses objetivos;
- Recrute as pessoas certas;
- Observe essas pessoas tentando executar as tarefas.

Tullis & Albert (2008) descrevem em detalhes as diversas métricas para se fazer uma avaliação de usabilidade:

- Métricas de desempenho;
- Métricas baseadas em Questões (*Issues*);
- Métricas de autorrelato;
- Métricas de comportamento e fisiológicas;
- Métricas combinadas e comparativas.

Entretanto, os diversos métodos citados até agora foram criados para avaliação genérica de usabilidade, sem uma preocupação específica em avaliar ambientes virtuais.

Nesse sentido, Bowman e coautores (2002) produziram um artigo que fez um levantamento abrangente de vários métodos para a avaliação de usabilidade em ambientes virtuais, listados a seguir:

- Percurso cognitivo;
- Avaliação Formativa;
- Avaliação heurística ou avaliação baseada em diretrizes realizada por especialista;
- Questionário pós-uso;

- Entrevista;
- Avaliação Somativa.

Este artigo compara duas abordagens distintas para a avaliação de usabilidade: a avaliação sequencial e a avaliação com *testbed*. A avaliação sequencial é feita no contexto de uma aplicação particular e fornece resultados tanto qualitativos como quantitativos. A avaliação com *testbed* é feita num contexto de avaliação genérico, e em geral busca resultados quantitativos.

Bowman e coautores (2005) produziram um livro onde consolidam os diversos métodos propostos para a avaliação de interfaces 3D. Eles destacam que um dos aspectos particulares de ambientes virtuais que deve ser avaliado é o senso de presença.

Nesse sentido, identificou-se na literatura o questionário IPQ, produzido pelo *igroup presence questionnaire* (<http://www.igroup.org/pq/ipq/index.php>), desenvolvido especificamente para avaliação de senso de presença.

O *Igroup Presence Questionnaire* (IPQ) é uma escala para medir o senso de presença experimentado num ambiente virtual. Ele foi construído usando-se uma grande quantidade de itens e dois levantamentos com aproximadamente 500 participantes. Ele foi originalmente desenvolvido em alemão, mas está também disponível em inglês e holandês.

A versão atual do IPQ tem três subescalas e um item geral que não pertence a uma subescala. As três escalas emergiram das análises dos principais componentes e podem ser vistas como fatores razoavelmente independentes. As subescalas são:

- Presença Espacial: o senso de estar fisicamente presente no ambiente virtual;
- Envolvimento: medindo a atenção devotada ao ambiente virtual e o envolvimento experimentado;
- Realismo Experimentado: medindo a experiência subjetiva de realismo no ambiente virtual.

A Avaliação Heurística foi um dos métodos para avaliação de usabilidade aplicados nesta dissertação. Ela é descrita por Nielsen (1993) como uma inspeção sistemática do design da interface do usuário, verificando o atendimento a princípios ou diretrizes de usabilidade.

De acordo com Nielsen (1993 p. 19), existem milhares de diretrizes para usabilidade, mas Nielsen (1993 p. 20) faz questão de destacar apenas 10 (dez):

- Diálogo simples e natural: não devem conter informações que sejam irrelevantes ou raramente necessárias;

- Falar a linguagem dos usuários: o diálogo deve ser expresso de forma clara, com palavras, frases e conceitos familiares aos usuários, ao invés do uso de termos do sistema;
- Minimizar a carga de memorização dos usuários: o usuário não deve ter de lembrar informação de uma parte do diálogo para outra;
- Consistência: usuários não devem ficar em dúvida se diferentes palavras, situações ou ações significam a mesma coisa;
- *Feedback*: o sistema deve sempre manter os usuários informados sobre o que está ocorrendo;
- Saídas claramente marcadas: usuários escolhem funções erradas com frequência e precisam uma “saída de emergência” claramente indicada, para poder sair com rapidez do estado indesejado;
- Atalhos: aceleradores, que não sejam percebidos pelo usuário novato, podem acelerar a interação do usuário avançado;
- Boas mensagens de erro: devem ser expressas em uma linguagem clara, indicar precisamente o problema e sugerir soluções;
- Prevenir erros: o design da interação deve ser cuidadoso para evitar o tanto quanto possível a ocorrência de erros;
- Sistema de ajuda e documentação: deve ser fácil de encontrar a informação procurada, focada na tarefa do usuário listar passos concretos a serem seguidos e não deve ser muito grande.

Essas 10 (dez) heurísticas de usabilidade propostas por Nielsen (1993) são intencionalmente genéricas, não abordando questões específicas de usabilidade voltadas para a avaliação de ambientes virtuais.

2.4. Avaliação de IHC em Ambientes Virtuais

Considerando que o foco deste estudo foi principalmente uma avaliação qualitativa de usabilidade de ambientes virtuais, foi escolhida uma abordagem de avaliação sequencial, na qual diversos métodos de avaliação são aplicados em sequência.

A abordagem de avaliação sequencial foi descrita por Gabbard e coautores (1999) num artigo que propõe uma metodologia para a avaliação de ambientes

virtuais (AVs) centrada no usuário (*User-Centered*), baseada na execução sequencial das seguintes etapas:

- Análise das tarefas do usuário;
- Avaliação baseada em diretrizes por um especialista (Avaliação Heurística);
- Avaliação formativa centrada no usuário;
- Avaliações somativas.

Este mesmo artigo detalha a sequência de métodos proposta:

- Análise das tarefas do usuário;
- Avaliação Heurística;
- Cenários de uso baseados em tarefas do usuário;
- Observação de uso com protocolo “Think-aloud”;
- Avaliadores sugerem melhorias para a interação;
- Avaliadores refinam os cenários de uso;
- Comparação do desempenho do usuário em designs alternativos.

Além de propor uma sequência de métodos, o artigo de Gabbard e coautores (1999) faz também a avaliação de usabilidade da aplicação *Dragon*, que se trata de um simulador de campo de batalha que faz uso de ambientes virtuais.

A avaliação de usabilidade desta dissertação foi desenvolvida com base nessa sequência de métodos proposta por Gabbard e coautores (1999), mas com modificações explicadas na seção 3.1, uma vez que esta sequência foi proposta para ser aplicada ao longo de todo o ciclo de desenvolvimento de um sistema. No nosso caso, o sistema avaliado era de um protótipo em estágio avançado de desenvolvimento, portanto não caberia a aplicação de todos os métodos.

Para realização de uma avaliação heurística, Bowman (2005) propôs diretrizes para o design de ambientes virtuais, mas elas de uma maneira geral são recomendações de alto nível no que se refere a usabilidade, dificultando a sua aplicação prática como diretrizes de avaliação. Por exemplo, como diretriz de design para locomoção (*travel*) ele recomenda “casar a técnica de locomoção com a aplicação” e “use uma combinação apropriada de técnica de locomoção, tela e dispositivos de entrada”. Para orientação (*wayfinding*), ele recomenda “casar a dica (*cue*) às habilidades do usuário” e “integre a locomoção com a orientação”. Para avaliação de interfaces 3D, ele recomenda “escolha uma abordagem de avaliação que atenda seus requisitos” e “use um largo espectro

de métricas”. Embora sejam diretrizes válidas, elas não oferecem muita informação para aplicá-las na prática.

Por outro lado, Gabbard (1997) produziu uma dissertação de Mestrado que propôs uma taxonomia e fez sugestões de usabilidade que podem ser usadas como diretrizes para uma Avaliação Heurística de AVs.

Em sua tese Gabbard coletou e sintetizou sugestões de usabilidade de várias fontes, incluindo de trabalhos de Nielsen e Bowman.

Gabbard propôs o que chama de um *framework* com quatro áreas principais:

- Usuário de AV e Tarefas de Usuário de AV;
- Interface do Usuário de AV – Mecanismos de Entrada;
- O Modelo Virtual;
- Interface do Usuário de AV – Componentes de Apresentação.

Além disso, ele subdividiu essas quatro áreas em sub-áreas, como ilustrado na Figura 1.

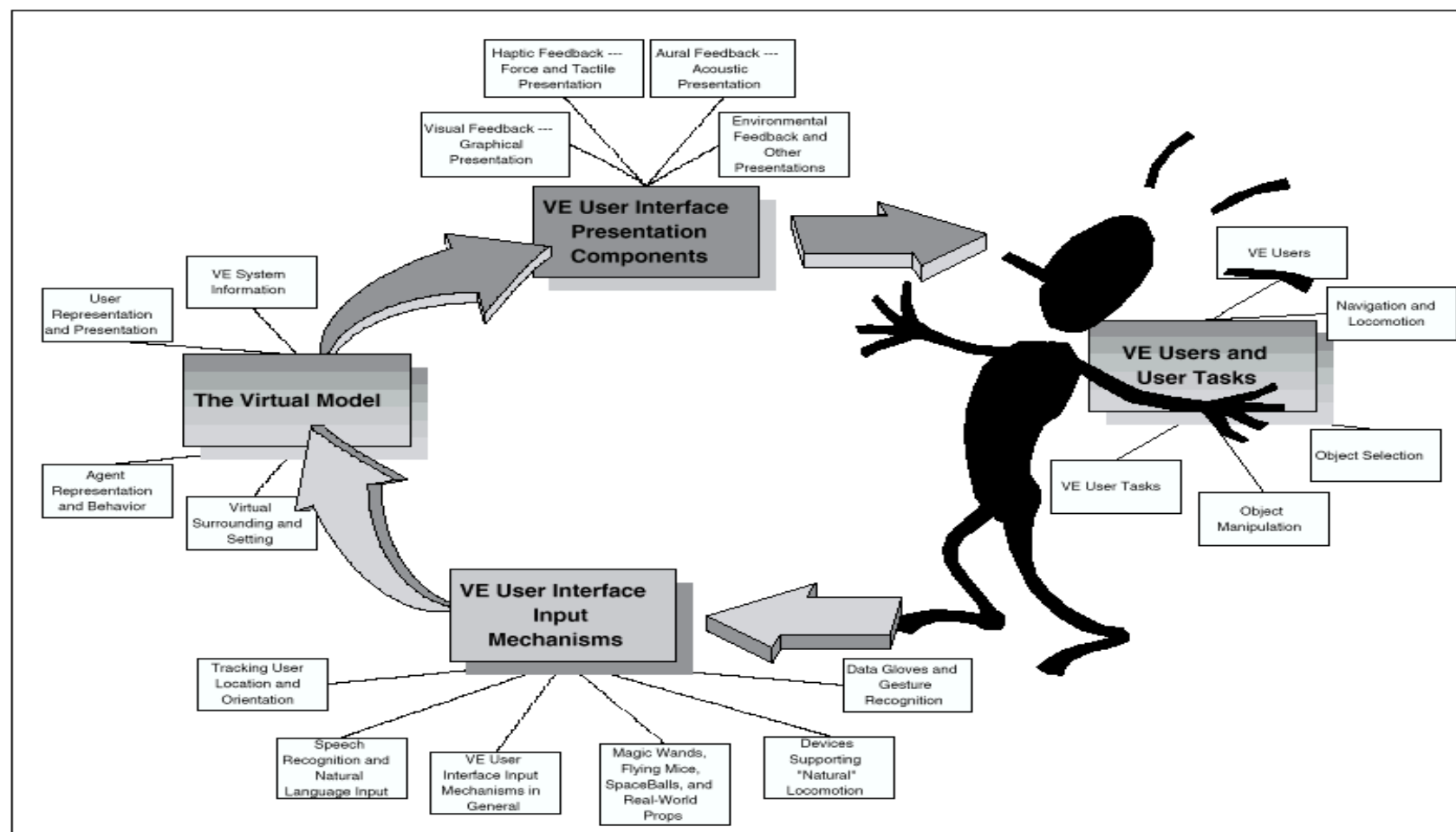
Segundo o próprio Gabbard (1997), a segmentação proposta na Figura 1, em quatro áreas principais, bem como o fluxo proposto, foram fortemente influenciados pela "teoria da ação" proposta por Norman (1990) e seus componentes correspondentes:

- Objetivos (*Goals*): Usuário de AV e Tarefas de Usuário de AV (*VE and User Tasks*);
- Execução: Interface do Usuário de AV - Mecanismos de Entrada (*VE User Interface Input Mechanisms*);
- O Mundo: O Modelo Virtual (*The Virtual Model*);
- Avaliação: Interface do Usuário de AV - Componentes de Apresentação (*VE User Interface Presentation Components*).

As mais de uma centena de sugestões, ou diretrizes de usabilidade, propostas em sua tese foram agrupadas nessas sub-áreas.

A Avaliação Heurística executada nesta dissertação foi aplicada com base nessas diretrizes propostas por Gabbard (1997).

No que se refere à avaliação de senso de presença em ambientes virtuais, Gentil (2008) produziu uma dissertação de Mestrado que faz uma avaliação de usabilidade da aplicação Second Life, aplicando o questionário de satisfação QUIS [Sharp et al. 2007], criado pelo laboratório de IHC da Universidade de Maryland, e o questionário IPQ (*igroup presence questionnaire*), para avaliação de senso de presença. Esta dissertação também aplicou o IPQ como método complementar ao proposto por Gabbard e coautores (1999).

Figura 1 - *Framework* de Características de Usabilidade em AV [Gabbard 1997]