

Avaliação de Interface em Aplicativo para Medição de Nível de Atividade Física Baseado no IPAQ da OMS

Arthur M. R. Alves, Victor B. L. de Paula, Yuri M. Ferreira,
Luis A. M. Mendes, Gabriella C. B. Costa, Matheus A. M. de Paula,
Luan S. Oliveira, Gabriela R. O. Venturini, Marcela G. Pinheiro,
João V. D. e Souza, Carlos H. S. Valente, Lorenzo J. B. Luz,
Antônio C. S. S. Neto, Lukas J. Wolf, Pedro F. S. Almeida,
Rodrigo A. S. Venetillo

Centro Federal de Educação Tecnológica de Minas Gerais
Leopoldina, Brasil

alves@aluno.cefetmg.br, vborgescontato@gmail.com, yurimoreiraferreira@hotmail.com,
luisaugusto@cefetmg.br, gabriella@cefetmg.br, matheusavila@cefetmg.br,
luan@cefetmg.br, doutoragabrielaventurini@gmail.com,
marcela.pinheiro@aluno.cefetmg.br, joao.souza@aluno.cefetmg.br,
carlos.valente@aluno.cefetmg.br, lorenzo.luz@aluno.cefetmg.br,
antonio.neto@aluno.cefetmg.br, lukas.wolf@aluno.cefetmg.br,
pedro.almeida@aluno.cefetmg.br, venetillo@aluno.cefetmg.br

Abstract

Due to globalization and digitalization, there has been a huge increase in the number of mobile applications being developed. This makes the user experience and usability of apps a topic of interest. This study evaluates an app based on the World Health Organization's International Physical Activity Questionnaire using Jakob Nielsen's heuristics as a framework, aiming to evaluate the final users perception of usability. The following ten criteria were considered: visibility of system statuses; correspondence between the real world and the system; user control and freedom; consistency and standards; error prevention; recognition rather than recall; flexibility, and efficiency of use; aesthetic and minimalist design; helping users to recognize, diagnose, and recover from errors; and finally, help and documentation. Therefore, the methodology used in this article made possible to evaluate the system - a scientific initiation's product - using Nielsen's heuristics as a baseline of study. In conclusion, this article makes recommendations for changes to the system en route to improve its usability as well the used methodology itself.

Keywords

Experience, Usability, Mobile Application, Heuristic

1 Introdução

Com a globalização e informatização através dos dispositivos eletrônicos, facilitou-se o acesso à informação, entretenimento e interação social entre as diversas pessoas e culturas de um contexto social, demográfico e geográfico. Esses dispositivos se tornaram ferramentas centrais no cotidiano das pessoas, transformando a forma com que trabalhamos, aprendemos e nos relacionamos [1]. Diante dos avanços da tecnologia, que acarretaram em um aumento no fluxo de informações que permeiam o dia a dia das pessoas, surge a necessidade de uma ferramenta científica que gerencie de forma mais organizada a integração Humano-Computador [2]. Essa

presença constante dos dispositivos faz com que a qualidade da interação seja aspecto crucial, influenciando diretamente o bem estar e produtividade dos usuários.

Ao tratarmos de um software e a sua interação com o homem, devemos levar em consideração como essa correlação se dá, sendo definida como *User Interface* (UI) - Interface de Usuário (pt-br) [3]. O *design* de UI influencia significativamente o engajamento de usuário sobre vários fatores, como: usabilidade, acessibilidade, estéticas visuais, interatividade e personalização [4]. Consequentemente, a garantia de que o resultado da interação do usuário seja satisfatória se torna fator determinante para uma inserção bem sucedida de um produto de software no mercado.

Portanto, é possível inferir que existe um resultado da ação de um usuário ao utilizar um sistema, definido como *User Experience* - Experiência de Usuário - (UX) [3]. O padrão ISO 9241-210 [5] define UX como "a percepção da pessoa e a resposta antecipada pelo uso de um produto, sistema ou serviço". Com isso, afirma-se que essa interação entre humano-computador resulta na experiência de uso do usuário. Além disso, as vendas e o sucesso de uma aplicação dependem intrinsecamente da UX e da UI, "Se a capa for interessante, as pessoas compram-na. Da mesma forma, se a tela da aplicação for confortável, o utilizador terá interesse em experimentá-lo. Só depois disso é que o utilizador verá as funções que lhe são oferecidas. A função ou a utilidade da aplicação também afeta as vendas de aplicações." [6].

Existem vários procedimentos e métricas que possibilitam mensurar e avaliar a qualidade de UX e UI implementadas em um software. Nesse contexto, as heurísticas de Nielsen [3], surgem como diretrizes para avaliar e aprimorar a usabilidade de interfaces, enumerando 10 heurísticas que são fundamentais para realizar essa análise: 1) visibilidade dos status do sistema; 2) correspondência entre o mundo real e o sistema; 3) controle e liberdade de usuário; 4) consistência e padrões; 5) prevenção de erro; 6) reconhecimento em vez de lembrança; 7) flexibilidade e eficiência de uso; 8) design

estético e minimalista; 9) ajudar os usuários a reconhecer, diagnosticar e se recuperar de erros; e, por fim, 10) ajuda e documentação [7].

Por fim, esse artigo possui como objetivo avaliar, de forma sistemática, a UX e UI de um aplicativo, produzido a partir de uma iniciação científica de uma instituição federal brasileira, utilizando as 10 heurísticas de Nielsen, proporcionando *feedback* e suporte para um processo de desenvolvimento incremental [8]. O sistema em questão possui em sua funcionalidade principal um formulário baseado no *International Physical Activity Questionnaire* (IPAQ) produzido pela Organização Mundial de Saúde (OMS), o questionário engloba quatro aspectos de atividade física: trabalho, transporte, serviços domésticos/jardinagem e lazer. Contando também com perguntas a respeito do tempo despendido desenvolvendo atividades sentado, a fim de identificar comportamentos sedentários [9].

O restante deste artigo está distribuído em Seções da seguinte forma: 2) apresenta o aplicativo sujeito a análise; 3) trabalhos correlatos que realizam avaliações de sistemas; 4) metodologia avaliativa utilizada a fim de analisar se o sistema - objeto de análise deste artigo - aplicou corretamente as heurísticas de Nielsen [3]; 5) resultados obtidos com a aplicação da metodologia; 6) discussão dos resultados obtidos apresentando soluções aos problemas encontrados; e, por fim, 7) conclui sobre os dados produzidos por este artigo e futuras aplicações.

2 Apresentação do Aplicativo

O aplicativo é fruto de uma iniciação científica realizada previamente por alguns dos autores. O sistema possui como objetivo principal de informatizar o processo de questionário de forma com que o usuário sinta pouca, ou nenhuma dificuldade em se adaptar - do papel impresso para a versão digital do questionário.

Como abordado brevemente na Seção 1 o aplicativo pretende informatizar o questionário *International Physical Activity Questionnaire* (IPAQ) produzido pela Organização Mundial de Saúde (OMS). O questionário foi desenvolvido como um instrumento para estabelecer uma ferramenta padronizada para medição de atividade física corriqueira, que pudesse ser adaptável a populações de diferentes culturas ao redor do mundo [10]. O IPAQ foi concebido para avaliar os níveis de atividade física habitual em pessoas de idade entre 15 a 69 anos de idade.

Além de proporcionar tais dados, o aplicativo permite que pesquisadores - pessoas que desejam agregar dados de um segmento da população e posteriormente analisá-los - possam fazer de forma informatizada a aplicação deste questionário. Com isso criam-se dois tipos de usuários no sistema: os pesquisadores e os participantes. O primeiro refere-se à pessoa interessada em levantar dados, e, posteriormente, realizar as devidas análises. O segundo são os que respondem ao questionário, são as pessoas que desejam obter informações baseadas nas suas respostas. No presente trabalho, apenas pessoas que utilizaram o aplicativo como o segundo tipo de usuário responderam o formulário de análise do sistema.

Para exemplificar o funcionamento do aplicativo algumas figuras foram selecionadas para demonstrar as diversas telas do sistema. A Figura 1 apresenta a tela inicial, login e de registro. A Figura 2 mostra três telas distintas. A primeira é a tela inicial após ser realizado o login no aplicativo, demonstrando as opções de realizar

o questionário ou visualizar os resultados; posteriormente há a tela dos resultados, que possibilita duas opções: visualizar o resultado no próprio aplicativo ou gerar o resultado em formato PDF; e, por fim, a última tela demonstra como são disponibilizados os resultados ao clicar na opção de visualizar os resultados no aplicativo. Ademais, por último, a Figura 3 demonstra o fluxo do questionário. Inicialmente a tela apresenta a opção de selecionar a sua localidade geográfica e a opção de responder como um visitante ou participante. A próxima tela é um informativo sobre a importância das respostas às questões; e a última apresenta um exemplo de como o questionário é distribuído - através de seções e perguntas com campos de preenchimento de dado.

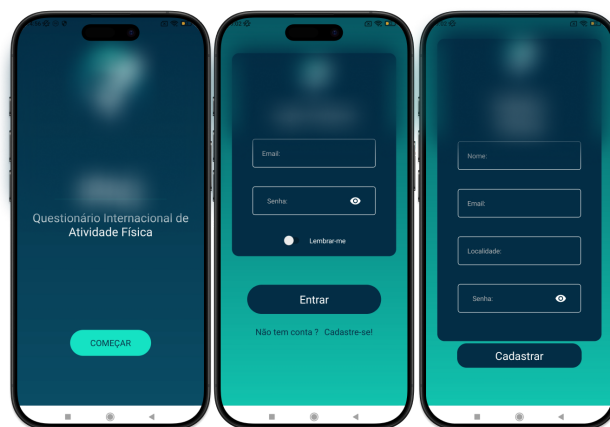


Figura 1: Respectivamente são as telas: inicial, login e registro

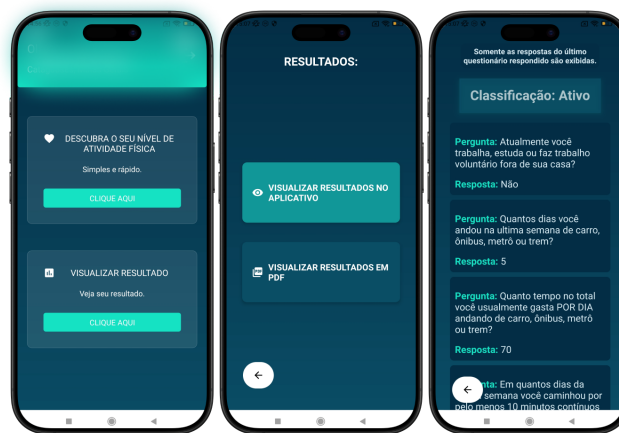


Figura 2: Respectivamente são as telas: menu de usuário, emissão de resultados e resultados

3 Trabalhos Correlatos

Para delimitar o cenário do estado da arte do tema de avaliações heurísticas de interações humano-computador, para aplicativos de celular, foram analisados os trabalhos a seguir, cada um oferecendo



Figura 3: Representação do fluxo do questionário

uma visão distinta sobre variados aspectos da avaliação da interação humano-computador com base na heurísticas de Nielsen.

Na monografia de Leite [11] a avaliação é feita a partir do Checklist para Avaliação da Usabilidade de Aplicativos para Celulares Touchscreen (MATCh) [12], desenvolvido pelo Grupo de Qualidade de Software (GQS) da Universidade Federal de Santa Catarina (UFSC), também baseado nas heurísticas de Nielsen e do System Usability Scale (SUS). É feita a avaliação do aplicativo *Spirit Fanfics* a partir de duas formas de avaliação, uma plataforma para autopublicação de livros. Após a análise dos resultados a autora conclui que "Embora o aplicativo apresente certas instabilidades em seu funcionamento, design e em questões de usabilidade, não é algo que definitivamente afetará a sua utilização. Todavia, ainda se torna necessário a melhoria para que usuários novatos não desistam da permanência no uso."

No trabalho de Canedo [13], foi feita a avaliação da usabilidade do sistema de jurisprudência do Conselho Administrativo de Defesa Econômica (CADE). Foram usadas as heurísticas estabelecidas por Da Costa et al. [14, 15] contendo 13 heurísticas e 196 sub-heurísticas, a avaliação das heurísticas foi feita por quatro especialistas da área. Foi identificado como problema que, considerando que os questionários eram extensos e demorados, os avaliadores podem ter se sentido cansados ou ter se distraído durante o processo. Podendo levar a erros nas avaliações finais das sub-heurísticas, como se esquecer de mencionar detalhes. Apesar das limitações o estudo foi capaz de demonstrar que o sistema de jurisprudência precisava de melhorias, como: campos de entrada de dados que não especificavam se eram de preenchimento obrigatório ou não e que ao completar uma tarefa ou conjunto de tarefas, o sistema não dava indicativos da possibilidade de iniciar uma nova ação.

Na tese de Scherer [16] é feita a avaliação de um software para *design* de interiores. As heurísticas de Nielsen foram adaptadas para o contexto da área de *design* de forma que para a construção das avaliações cada heurística pudesse ser avaliada visualmente na forma de uma captura de tela de uma funcionalidade da aplicação. O SUS foi utilizado como segunda forma de avaliação, onde não foram necessárias adaptações. Ambas as avaliações foram aplicadas no software *Sweet Home 3D*. As avaliações foram aplicadas por

cinco especialistas na área de integração humano-computador, dois deles possuíam formação na área de arquitetura e urbanismo. Foi aplicado um minicurso para capacitar os avaliadores ao manuseio do programa, em seguida, entrou-se em contato com algumas instituições educacionais para o recrutamento de professores, alunos e profissionais da área que pudessem servir como usuários avaliadores. 32 Alunos foram confirmados. Os resultados evidenciaram a importância de combinar diferentes métodos de avaliação para identificar uma maior variedade de problemas em interfaces. Na avaliação heurística, o principal problema foi relacionado à estética e design minimalista, não detectado por outros avaliadores ou no teste de usabilidade. Já o teste de usabilidade revelou falhas como o mau funcionamento do zoom, não percebidas pelos especialistas. O trabalho contribuiu com heurísticas específicas para design de interiores e destacou a eficácia do questionário SUS para avaliar e classificar a qualidade da interface do Sweet Home 3D, identificando componentes que precisam de melhorias e captando a opinião dos usuários.

Os trabalhos apresentados tem como foco avaliar aplicações amplamente difundidas e voltadas para um público-alvo específico de usuários finais, enquanto o presente estudo foca na análise de um software desenvolvido como produto de uma iniciação científica, destacando-se por ser um projeto acadêmico realizado pelos próprios autores na instituição, como ferramenta de coleta de dados para pesquisas na área de saúde. Também é possível identificar a diferença no volume e de especialização dos avaliadores, levando em conta que 153 alunos avaliaram o aplicativo baseado no IPAQ, desempenhando o papel de usuário final, portanto, A avaliação se caracteriza como uma percepção de usabilidade com base nas heurísticas de Nielsen. O presente trabalho se destaca pelo foco em um aplicativo específico para a área de saúde e por apresentar uma amostra grande de usuários.

4 Metodologia

4.1 Criação do Questionário Avaliativo

Um questionário avaliativo foi criado baseado nas 10 heurísticas de Nielsen[3] com o intuito de avaliar a UI e UX do aplicativo em questão a fim de orientar o *design* e desenvolvimento futuro das interfaces - desenvolvimento incremental [7]. Portanto, para levantar os dados em questão - objetos de análise deste artigo - utilizou-se o *feedback* de usuários que testaram a primeira versão do sistema.

Consequentemente, com a finalidade de obter uma avaliação minuciosa e também heurística do aplicativo, traduziu-se cada uma das dez heurísticas de Nielsen[3] no formato de perguntas que pudessem proporcionar o sentido mais fiel possível de cada uma delas. Dessa forma, em cada tópico, foi decidido aplicar perguntas na forma objetiva e que abordassem de forma direta e sem dúvida o tópico da heurística[17].

Evitando a limitação de apenas perguntar ao usuário se ele concorda ou discorda, a abordagem escolhida foi utilizar a escala Likert. Os itens da escala Likert são usados para ranquear o quanto um usuário concorda ou discorda, geralmente em uma escala de cinco ou sete pontos. Sendo 1 "discordo fortemente" e 5 ou 7 "concordo fortemente", enquanto 3 (no caso de total 5) ou 4 (no caso de total 7) pontos que representam um posicionamento de neutralidade[18].

A opção pela escolha do usuário final como avaliador

A opção de classificação de respostas escolhida neste questionário avaliativo foi a de cinco. Dessa forma, em todas as perguntas objetivas foram dispostas cinco opções de seleção de resposta de seleção única, sendo elas: discordo, discordo parcialmente, indiferente, concordo parcialmente e concordo, e seus sinônimos [19]. Respectivamente, seguindo a escala Likert, variando dos valores 1 à 5. Valendo ressaltar que cada valor semântico foi transcrito para que fizesse sentido com a pergunta realizada, vide Tabela 1. Nela estão expostas as descrições de cada um dos tópicos da heurística de Nielsen; e a correlação das perguntas formuladas. As colunas estão distribuídas no seguinte sentido: o atributo "Id" é o valor de identificação da heurística - que será utilizado nas diversas tabelas deste artigo; "Heurística", o nome de cada uma delas; e, por fim, "Perguntas", texto utilizado na pergunta.

Com a finalização da criação das perguntas objetivas, também foram criadas perguntas com respostas na modalidade discursiva. A utilização desse tipo de questão está diretamente relacionada à possibilidade de explicação e síntese textual do pensamento dos participantes - avaliadores do sistema, UI e UX. Por fim, elas foram objetos de análise na Seção 6 a fim de explorar adequadamente os resultados da avaliação.

Por fim, o restante desta Seção explora cada uma das heurísticas a fim de demonstrar os motivos relacionados que motivaram a seleção de texto à cada uma delas para descreve-las através de perguntas.

4.1.1 Visibilidade dos status do sistema. "O sistema deve informar continuamente o usuário sobre o que está fazendo e como está interpretando a entrada do usuário. O feedback deve não esperar até que uma situação de erro ocorra: o sistema deve também fornecer feedback positivo e deve fornecer feedback parcial à medida que a informação se torna disponível [...]. O feedback do sistema não deve ser expresso de forma abstrata em termos gerais, mas deve formular e reformular a entrada do usuário para indicar o que está sendo feito com isso." [3]. Dessa forma o sistema deve, a partir de qualquer ação tomada pelo usuário, fornecer um feedback com o intuito de informar o estado do sistema da forma mais clara possível.

4.1.2 Correspondência entre o mundo real e o sistema. "Como parte do princípio geral de falar a língua dos utilizadores, deve-se tomar cuidado para não usar palavras com significados fora do padrão, a menos, é claro, que ela signifique algo na comunidade de usuários e que não seja padrão na população em geral" [3]. Assim o sistema deve apresentar correspondência ao mundo real, com linguagem clara e objetiva. No caso do objeto de estudo, o aplicativo em questão deve proporcionar ao usuário a mesma experiência ao preencher um formulário da forma usual, impresso, utilizando as mesmas terminologias e metodologias de aplicação.

4.1.3 Controle e liberdade de usuário. "Os usuários não gostam de se sentir presos ao computador. A fim de aumentar a sensação do usuário de estar no controle do diálogo, o sistema deve oferecer ao usuário uma maneira fácil de sair de tantas situações possível [...]. Os usuários aprendem rapidamente a confiar na existência do desfazer, então deve ser disponibilizado de forma generalizada em todo o sistema como um comando genérico que desfaz qualquer mudança de estado em vez de ser restrito a desfazer apenas uma categoria especial de ações do usuário. Dado que os recursos de

desfazer e escapar estão geralmente disponíveis, os usuários sentem-se encorajados a confiar na aprendizagem exploratória, uma vez que podem sempre experimentar opções desconhecidas, confiando em sua capacidade de sair de qualquer problema sem efeitos nocivos" [3].

O usuário deve sentir que está no controle do aplicativo, ou seja, ao clicar em um botão, como retornar à uma tela anterior, o aplicativo deve responder de forma imediata a ação e retornar ao estado anterior. O aplicativo deve ainda evitar erros no sistema, como fechamento inesperado da interface gráfica. Ademais, a partir de uma execução de espera do aplicativo deve ser possibilitado ao usuário cancelar a ação e voltar a um estado anterior.

4.1.4 Consistência e padrões. "A consistência é um dos princípios de usabilidade mais básicos. Se os usuários sabem que o mesmo comando ou a mesma ação sempre terá o mesmo efeito, eles se sentirão mais confiantes no uso do sistema, e serão incentivados a experimentar estratégias de aprendizagem exploratória porque já terão parte do conhecimento necessário para operar novas partes do sistema" [3, 20]. Ademais, as cores e modelos dos objetos da interface devem permanecer os mesmos durante todas as telas da aplicação. Portanto, para as diversas telas da aplicação, deve ser estruturado um padrão, uma normatização, que deve ser seguida em todos os pontos do aplicativo.

4.1.5 Prevenção de erro. "Erros com consequências especialmente graves também podem ser reduzidos em frequência, pedindo aos usuários que confirmem que "realmente, realmente quero dizer isso" antes de prosseguir com as ações perigosas. Não devemos usar diálogos de confirmação com tanta frequência, porém, a resposta do usuário torna-se automática. Se uma longa sequência de ações for realizado com tanta frequência que é experimentado como uma unidade, os usuários correm o risco de cometer um "erro de captura", e conclui: se algum dia precisarem desviar-se da sequência: porque estão tão acostumados a seguir em frente de certa forma, eles podem continuar e emitir o clique fatal no OK antes mesmo de lerem a mensagem de aviso" [21].

Portanto, a questão de evitar um erro por parte do usuário - uma interface que evite uma ação errônea do mesmo - como a correta descrição de uma mensagem informativa que possibilite-o a compreender o que está sendo exposto.

4.1.6 Reconhecimento em vez de lembrança. "Os computadores são muito bons em lembrar coisas com muita precisão, então eles devem assumir a carga de memória do usuário tanto quanto possível. Em geral, as pessoas têm muito mais facilidade em reconhecer algo que lhes é mostrado do que eles têm em ter que recuperar as mesmas informações da memória sem ajuda" [3]. Sob essa perspectiva, uma padronização e correta visibilidade das diversas funções do sistema, necessita de uma interface que faça o usuário reconhecer certos padrões. Um grande exemplo que ocorre nos diversos sistemas é a utilização de uma imagem de disquete para representar a ação de persistência de dados - salvamento -, dado que em um primeiro momento o usuário não pensa nisso de forma ativa, porém, ao ser mostrado o ícone, a associação e a lembrança de seu significado surge.

4.1.7 Flexibilidade e eficiência de uso. "Embora seja possível operar uma interface de usuário com conhecimento de apenas algumas

Tabela 1: Perguntas objetivas para avaliar o aplicativo baseado nas 10 heurísticas de Nielsen

ID	Heurísticas	Perguntas
1	visibilidade dos status do sistema	Durante o uso do sistema, você se sentiu informado sobre o que estava acontecendo?
2	correspondência entre o mundo real e o sistema	As informações e termos usados no sistema eram claros e familiares para você?
3	controle e liberdade de usuário	Quando você cometeu um erro, foi fácil corrigi-lo?
4	consistência e padrões	A interface do sistema parecia consistente?
5	prevenção de erro	O sistema ajudou você a evitar erros?
6	reconhecimento em vez de lembrança;	Foi fácil encontrar as informações ou funcionalidades que você precisava, ou você precisou consultar o material disponível para usá-las?
7	flexibilidade e eficiência de uso;	Você sentiu que o sistema permitiu que você realizasse as tarefas de maneira rápida e eficiente?
8	design estético e minimalista	A interface parecia limpa e organizada?
9	ajudar os usuários a reconhecer, diagnosticar e se recuperar de erros;	Quando algo deu errado, as mensagens de erro foram úteis?
10	ajuda e documentação.	Se você precisasse de ajuda, o sistema ofereceu suporte suficiente?

regras gerais, também deveria ser possível para o usuário experiente executar operações usadas com frequência especialmente rápido, usando atalhos de diálogo. Aceleradores típicos incluem abreviações, tendo teclas de função ou teclas de comando que empacotar um comando inteiro com um único toque de tecla, clicando duas vezes em um objeto para realizar a operação mais comum nele, e ter botões disponíveis para acessar funções importantes diretamente das partes do diálogo onde podem ser mais frequentemente necessário"[3]. O sistema deve proporcionar agilidade ao usuário em suas ações. Uma exemplificação desse tipo de eficiência é autocompletar informações baseadas em alguma ação anterior do usuário. Por exemplo, nos mais variados sistemas que implementam uma tela de cadastro que é necessário o fornecimento de um endereço é possível inserir um número de Código de Endereçamento Postal (CEP) que é vinculado a um estado, cidade, bairro e rua, preenchendo automaticamente esses campos.

4.1.8 Design estético e minimalista. "As interfaces do usuário devem ser simplificadas tanto quanto possível, uma vez que cada recurso ou item de informação adicional em uma tela é mais uma coisa para aprender, mais uma coisa para possivelmente interpretar mal, e mais uma coisa para pesquisar"[3]. Dessa forma, no geral, a aplicação deve ser funcional e agradável ao usuário, portanto deve-se levar em consideração a experiência que o usuário terá ao utilizar o sistema. Além disso, as informações dispostas ao usuário devem ser apenas as necessárias para a utilização do sistema, devendo evitar quaisquer informações que não sejam alheias a ele.

4.1.9 Ajudar os usuários a reconhecer, diagnosticar e se recuperar de erros. "Situações de erro são críticas para a usabilidade por dois motivos: Primeiro, por definição, eles representam situações em que o usuário está com problemas e potencialmente será incapaz de usar o sistema, para alcançar o desejado propósito final. Em segundo lugar, eles oferecem oportunidades para ajudar o usuário a entender melhor o sistema"[22]. Portanto, deve abordar unicamente a correta descrição e exposição dos erros que ocorrem no sistema

ao usuário, de forma que seja possível o mesmo realizar ações que o possibilite a voltar a um estado anterior ao erro e fazer as devidas correções necessárias, se possível.

4.1.10 Ajuda e documentação. "Embora seja preferível que um sistema seja tão fácil de que ajuda ou documentação não é necessária para complementar a interface do usuário em si, porém esse objetivo nem sempre pode ser alcançado. Exceto para sistemas como caixas eletrônicos onde verdadeiros *walk-up-and-use* em que a usabilidade é necessária, a maioria das interfaces de usuário tem número suficiente de recursos para garantir um manual e possivelmente um sistema de ajuda. Também, usuários regulares de um sistema podem querer documentação que lhes permitam adquirir níveis mais elevados de conhecimento." [3] Em todo sistema um ambiente de ajuda e sua documentação deve ser fornecida ao usuário a fim de servir como suporte.

4.2 Método de Avaliação

Dois alunos do curso de Engenharia de Computação analisaram os dados resultantes da aplicação do questionário formulado por este artigo. Alguns métodos de avaliação foram definidos com a finalidade de mensurar os erros encontrados no sistema. As duas métricas utilizadas que possibilitaram essa análise são: **aplicação** e **qualidade**. A primeira é a frequência - traduzida em valores de porcentagem referentes ao total de respostas - e a outra será a média ponderada - calculada através da multiplicação pelos valores respectivos, de 1 a 5, e dividido pela soma das pontuações da escala Likert que totalizam 15 pontos, obtendo assim uma média ponderada. Respectivamente, as Subseções 4.2.1 e 4.2.2 exemplificam cada uma delas.

4.2.1 Aplicação. Esta métrica nos deu o indicativo de que a heurística foi violada. Definiu-se que, para calcular a porcentagem, se dará pela seguinte equação:

$$f(x) = \frac{x * 100\%}{\alpha} \quad (1)$$

Tabela 2: Resultado da Aplicação e Qualidade das heurísticas

ID	Aplicação(%)	Qualidade
1	60,13	4,44
2	73,20	4,63
3	31,37	3,75
4	54,25	4,35
5	30,07	3,60
6	56,21	4,35
7	57,52	4,28
8	73,86	4,53
9	36,60	3,79
10	44,44	3,95

, sendo α o valor total de respostas, que no caso desta pesquisa foram 153, e x a frequência com que uma resposta aparece.

Com a utilização dessa equação temos a porcentagem da frequência de cada uma das respostas da escala Likert, ou seja, em qualquer questão do questionário teremos as frequências das respostas, valores de um à cinco de cada uma delas. A partir destes valores, a fim de definir se há algo errado com a aplicação heurística, utilizou-se a resposta de maior valor da escala, Likert - valor cinco "concordo fortemente" - como limiar de reprovação de 60%. Com isso sugere-se que se ela não agradou pelo menos 60% do público com a melhor avaliação, então ela apresentará algum tipo de problema.

4.2.2 Qualidade. Já na segunda métrica, através da média ponderada, será possível identificar se o aplicativo obteve uma boa avaliação na heurística. Sendo calculada da seguinte forma:

$$\frac{\sum_{i=1}^5 ix_i}{\alpha} \quad (2)$$

, onde α é o número de respostas totais da pesquisa e x_i é a frequência da resposta i . A partir desse valor é possível mensurar a nota de uma avaliação da aplicação da heurística no aplicativo. Para reprovar a **qualidade de aplicação** da heurística, a pontuação deverá ser inferior a 3.

5 Resultados

Conforme o método de avaliação, a partir da verificação das aplicações das heurísticas no sistema, com os dados da Tabela 2, podemos verificar que não foram corretamente aplicadas: 3) controle e liberdade de usuário; 4) consistência e padrões; 5) prevenção de erro; 6) reconhecimento em vez de lembrança; 7) flexibilidade e eficiência de uso; 9) ajudar os usuários a reconhecer, diagnosticar e se recuperar de erros; e 10) ajuda e documentação. Já pela métrica da qualidade, nenhuma aplicação da heurística foi reprovada.

6 Discussão dos Resultados

Conforme os dados da Tabela 2, 7 das heurísticas abordadas retornaram avaliações de que o sistema não proporcionou uma experiência satisfatória aos usuários. A partir disso, é necessário que se analise as questões abertas, a fim de identificar quais foram os problemas. A análise dessas questões ocorreu conforme a métrica de falha na aplicação da heurística.

6.1 Controle e liberdade de usuário

Com apenas 58,16% dos participantes avaliando a heurística com respostas positivas, 28,10% neutras e 13,74% negativas, uma avaliação das questões discursivas se tornou necessária. Ao analisar essas respostas, foram identificados alguns erros críticos do sistema: travamentos e reinício, que causavam problemas para os usuários ao cometerem erros no durante o uso.

6.2 Consistência e padrões

Nesta heurística, apenas 86,27% das avaliações deram notas positivas, 9,15% neutras e apenas 4,58% negativas. Com as respostas das questões apenas uma inconsistência foi observada: tamanho da fonte diferente em partes similares do sistema, como demonstra a Figura 4.



Figura 4: Exemplo de inconsistência nos tamanhos de fonte.

6.3 Prevenção de erro

A única resposta que fazia referência a heurística foi sobre a validação de dados preenchidos em alguns campos, fazendo com que o usuário fosse forçado a voltar a telas anteriores para corrigir a resposta, porém foi identificado que era necessário alertar os usuários sobre a forma com que alguns dados deveriam ser inseridos, como demonstra a Figura 5.

6.4 Reconhecimento em vez de lembrança

Para esta heurística em específico não haviam questões abertas relacionadas e não foi possível identificar o possível erro observado pelos participantes da pesquisa.

Tabela 3: Soluções para as heurísticas que falharam na aplicabilidade

ID	Heurísticas	Soluções
3	controle e liberdade de usuário	Realizar testes no sistema em busca dos erros que faziam o aplicativo parar de funcionar.
4	consistência e padrões	Padronizar o tamanho da fonte nas diversas telas do sistema.
5	prevenção de erro	Formatação correta dos tipos dos campos de dados e fornecimento de informações suficientes aos usuários de como deve ser feito esse preenchimento.
6	reconhecimento em vez de lembrança	x
7	flexibilidade e eficiência de uso	Implementar script para preencher os campos e formatar conforme a inserção de caracteres.
9	ajudar os usuários a reconhecer, diagnosticar e se recuperar de erros	Implementar mensagens de erros descritivos.
10	ajuda e documentação	Implementação de uma documentação externa e possibilitar o redirecionamento para ela de dentro do sistema.

6.5 Flexibilidade e eficiência de uso

No geral, as respostas discursivas possuíram uma classificação positiva em relação à heurística, o que explica sua frequência na melhor nota da escala Likert, 83,66% positivas e apenas 10,45% negativas, com o restante sendo neutras. Apesar disso, foi elicitado a necessidade de implementar um autopreenchimento de caracteres especiais para formatar a entrada do usuário.



Figura 5: Exemplo de entrada incorreta de dados seguida de como a entrada deveria ser feita.

Mostrar foto do campo de dado que deveria ser do formato de data e a fonte com tamanho diferente.

6.6 Ajudar os usuários a reconhecer, diagnosticar e se recuperar de erros

As respostas negativas não explicavam o motivo de não terem conseguido reconhecer, diagnosticar ou se recuperar de erros. Entretanto algumas informam que não foi observado informações sobre os erros que estavam ocorrendo.

6.7 Ajuda e documentação

Quando foi feita a pesquisa sobre o aplicativo haviam pessoas que tiravam dúvidas sobre o sistema caso os participantes estivessem com dúvida. Dessa forma, nas questões abertas, muitos usuários relataram a necessidade de ajuda externa pois não encontravam a resposta no sistema.

A partir disso, para cada uma das heurísticas é apresentada uma possível solução. Na Tabela 3 é possível observar cada um desses tópicos.

7 Conclusão

Conforme explicitado na Seção 1, a tecnologia está cada vez mais embutida no cotidiano dos humanos, e é necessário um método de análise sobre essa interação humano-computador. Nielsen cria as dez heurísticas que possibilitam analisar se um sistema apresenta uma boa interação humano-computador (UI) e como essa experiência se dá (UX).

Ao se tratar do questionário, é possível observar que cada pergunta foi formulada para explicitar adequadamente cada uma das heurísticas de forma singular e objetiva. Ademais, a fim de realizar uma análise adequada sobre os dados, a estruturação das respostas utilizando a escala Likert possibilitou fazer um julgamento sobre a aplicabilidade e qualidade do sistema no que tange a cada uma das heurísticas. Além disso, as questões de formato discursivo possibilitaram explicitar exatamente os tópicos que precisam ser melhorados no sistema. Entretanto, pode-se observar que na heurística “reconhecimento em vez de lembrança” houve necessidade de relacionamento a uma pergunta discursiva, pois a falta de uma foi o fator que impossibilitou a verificação na falha da aplicabilidade.

Conforme na Seção 6 foi possível analisar as respostas das questões discursivas e apresentar possíveis soluções ao sistema a fim de resolver os problemas elicitados.

Portanto, através da análise realizada, foi possível inferir quais os pontos a serem trabalhados na aplicação, identificar problemas nas aplicações das heurísticas através do questionário desenvolvido para análise do sistema, tal como sugerir possíveis soluções aos problemas encontrados na aplicação. Ademais, os resultados deste trabalho torna possível o aprimoramento do questionário avaliativo utilizado, vide Seção 4, para futuras avaliações se sistemas oriundas de futuras pesquisas e trabalhos.

Referências

- [1] Marco Fasoli. 2021. The overuse of digital technologies: Human weaknesses, design strategies and ethical concerns. *Philos. Technol.* 34, 4 (2021), 1409–1427.
- [2] David M Levy. 2005. To grow in wisdom: Vannevar bush, information overload, and the life of leisure. In *Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries*. ACM, New York, NY, USA.
- [3] Jakob Nielsen. 1995. *Usability Engineering*. Morgan Kaufmann, Oxford, England.
- [4] Ar Poorva Priyadarshini. 2024. The Impact of User Interface Design on User Engagement. *International Journal of Engineering Research Technology (IJERT)* 13 (2024), Issue 3.
- [5] International Organization for Standardization. 2019. *ISO 9241-210:2019: Ergonomics of human–system interaction – Part 210: Human-centred design for interactive systems*. International Organization for Standardization, Geneva.
- [6] Muhammad Aulia Taufik Pratama and Ahmad Taufiq Cahyadi. 2020. Effect of User Interface and User Experience on application sales. *IOP Conf. Ser. Mater. Sci. Eng.* 879, 1 (2020), 012133.
- [7] Mary Beth Rosson and John M Carroll. 2014. *Usability engineering: Scenario-based development of human-computer interaction*. Morgan Kaufmann.
- [8] Suzanne Kieffer, Aissa Ghouti, and Benoit Macq. 2017. The agile UX development lifecycle: Combining formative usability and agile methods. In *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*. Hawaii International Conference on System Sciences.
- [9] Maria Hagströmer, Pekka Oja, and Michael Sjöström. 2006. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutr.* 9, 6 (2006), 755–762.
- [10] Youngdeok Kim, Ilhyeok Park, and Minsoo Kang. 2013. Convergent validity of the international physical activity questionnaire (IPAQ): meta-analysis. *Public Health Nutr.* 16, 3 (2013), 440–452.
- [11] Joana Morais Santos Leite. 2023. *Avaliação de usabilidade em plataformas de autopublicação de livros: uma análise do Spirit Fanfics*. IFBA, Departamento de Ensino Técnico, Seabra, BA.
- [12] Luiz Henrique A Salazar, Thaísa Lacerda, Juliane Vargas Nunes, and Christiane Gresse von Wangenheim. 2013. A systematic literature review on usability heuristics for mobile phones. *Int. J. Mob. Hum. Comput. Interact.* 5, 2 (2013), 50–61.
- [13] Edna Dias Canedo, Ana Paula Morais do Vale, Rafael Leite Patrão, Leomar Camargo de Souza, Rogério Machado Gravina, Vinicius Eloy dos Reis, Felipe Alberto Moreira Dias, Fábio Lúcio Lopes Mendonça, and Rafael T. de Sousa. 2021. Usability Assessment of a Jurisprudence System. In *HCI in Business, Government and Organizations*, Fiona Fui-Hoon Nah and Keng Siau (Eds.). Springer International Publishing, Cham, 482–499.
- [14] Ruyther Parente da Costa and Edna Dias Canedo. 2019. A Set of Usability Heuristics for Mobile Applications. In *Human-Computer Interaction. Perspectives on Design*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 180–193.
- [15] Ruyther Parente Da Costa, Edna Dias Canedo, Rafael Timóteo De Sousa, Robson De Oliveira Albuquerque, and Luis Javier García Villalba. 2019. Set of Usability Heuristics for Quality Assessment of Mobile Applications on Smartphones. *IEEE Access* 7 (2019), 116145–116161. <https://doi.org/10.1109/ACCESS.2019.2910778>
- [16] Noemi Pereira Scherer. 2018. Avaliação heurística e teste de usabilidade para softwares de design de interiores. , 53 pages. <http://repositorio.utfpr.edu.br/jspui/handle/1/6009>
- [17] Setia Hermawati and Glyn Lawson. 2016. Establishing usability heuristics for heuristics evaluation in a specific domain: Is there a consensus? *Appl. Ergon.* 56 (2016), 34–51.
- [18] John Robinson. 2023. Likert Scale. In *Encyclopedia of Quality of Life and Well-Being Research*. Springer International Publishing, Cham, 3917–3918.
- [19] Gail M Sullivan and Anthony R Artino, Jr. 2013. Analyzing and interpreting data from likert- type scales. *Journal of graduate medical education* 5, 4 (2013), 541–542.
- [20] Charles Lewis, David C Hair, and Victoria Schoenberg. 1989. Generalization, consistency, and control. *SIGCHI Bull.* 20, SI (1989), 1–5.
- [21] Donald A Norman. 1983. Design principles for human-computer interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems - CHI '83*. ACM Press, New York, New York, USA.
- [22] Michael Frese, Felix Brodbeck, Torsten Heinbokel, Christina Mooser, Erik Schleiffenbaum, and Petra Thiemann. 1991. Errors in training computer skills: On the positive function of errors. *Hum.-Comput. Interact.* 6, 1 (1991), 77–93.