

Hidrate Spark: Avaliando um Sistema Ubíquo para Motivar a Ingestão de Água

José Cezar de Souza Filho, Marcos Randel Freitas Brito, Antonio Raimundo Rocha Mendonça,
Marcos Vitor Martins, Andréia Libório Sampaio

Universidade Federal do Ceará – Campus Quixadá
Quixadá, Brasil

{cezarbx, randelfreitas, antonioraimundo007, markstation22, andreia.ufc}@gmail.com

RESUMO

O objetivo do presente artigo é avaliar a usabilidade e a experiência do usuário do aplicativo ubíquo *Hidrate Spark*, desenvolvido para apoiar a ingestão diária de água. Para isso, um conjunto de métodos foi utilizado, a fim de enfrentar o desafio de avaliar a qualidade de uso de sistemas ubíquos. Os resultados mostram os problemas que podem prejudicar a interação do usuário com o aplicativo.

Palavras-chave

Sistemas Ubíquos; Usabilidade; Experiência do Usuário; Heurísticas Ubíquas; Maltu; Teste de Usabilidade.

INTRODUÇÃO

Considerando que o tema da competição de avaliação foi avaliar como a tecnologia pode apoiar as pessoas nos cuidados diários com a saúde, escolhemos avaliar a usabilidade do sistema *Hidrate Spark*, o qual possui integração com uma garrafa inteligente. O aplicativo e a garrafa compõem um sistema ubíquo, desenvolvido para auxiliar os usuários no seu monitoramento diário de ingestão de água, o que nos motivou para a escolha do aplicativo, visto que avaliar a qualidade de uso de um sistema ubíquo não é uma tarefa trivial, considerando que esse domínio de sistema interativo apresenta características específicas em comparação com sistemas tradicionais (e.g., desktop e web), tais como: sensibilidade ao contexto, mobilidade e transparência [3]. Além disso, o aplicativo foi escolhido porque possui versões gratuitas disponíveis nas lojas virtuais da plataforma Android (*Play Store*) e iOS (*App Store*), e por disponibilizar um fórum de discussão¹ para os usuários relatarem os problemas encontrados.

O número de usuários (2517 somente na plataforma Android) mostra que o aplicativo é relevante para determinadas pessoas e favorece a aplicação de alguns

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Copyright 2018 SBC.

IHC 2018, Anais Estendidos do XVII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais
Outubro 22–26, 2018, Belém, Brasil
Competição de Avaliação - Graduação

¹ <https://hidratespark.com/pages/reviews>

métodos de avaliação. Através da escala de 1 a 5 estrelas, os usuários avaliaram o aplicativo com valor 3.6 na *Play Store*² e 3.0 na *App Store*³, e a garrafa recebeu avaliação 3.2 na Amazon⁴ (*e-commerce* onde a garrafa pode ser adquirida). Esses dados foram obtidos no dia 10 de Junho de 2018 e comprovam a necessidade de uma avaliação, a fim de melhorar a qualidade de uso dos aplicativos.

É importante ressaltar que os aplicativos disponíveis para ambos os sistemas operacionais possuem as mesmas funcionalidades, permitindo a identificação de problemas de usabilidade existentes nos dois aplicativos ou apenas em um deles. Essa relação entre os aplicativos é comprovada pelo seguinte comentário de um usuário:

“Adoro essa garrafa e uso desde quando tinha um iPhone. Porém estou tendo problemas no Android desde que troquei de celular. Hora não sincroniza, hora dá crash ao abrir o app. Gostaria de uma ajuda pra solucionar isso.” (Fonte: *Play Store*).

Considerando essas questões, e considerando que a forma de interação com esses sistemas muda em relação aos sistemas convencionais [9] o objetivo deste artigo é avaliar a usabilidade e a experiência do usuário (UX, do inglês *User eXperience*) do *Hidrate Spark*.

Para avaliar a usabilidade desses sistemas, é preciso considerar características específicas em comparação com sistemas tradicionais, tais como: sensibilidade ao contexto, mobilidade, calma e transparência [3]. Alguns trabalhos investigaram como os métodos tradicionais de avaliação de usabilidade estão sendo usados para avaliar sistemas ubíquos. Bonifácio *et al.* [2] mostraram possibilidades para adaptar métodos de inspeção para avaliação da usabilidade de sistemas móveis e web, e dentre eles, as Heurísticas de Nielsen são as mais usadas. Rocha *et al.* [10] propuseram Heurísticas Ubíquas para avaliar a usabilidade de sistemas ubíquos, denominadas HUBIS, fundamentadas no método de Nielsen.

Para avaliar a UX utilizamos o MALTU [5], um modelo que analisa os comentários sobre usabilidade e UX a partir de postagens em um sistema social. Em relação à

² <https://play.google.com/store>

³ <https://www.apple.com/itunes/>

⁴ <https://www.amazon.com/>

usabilidade, usamos o MALTU, além do Teste de Usabilidade [1] e a Avaliação Heurística com regras adaptadas para Sistemas Ubíquos [6,10]. O Teste de Usabilidade também foi usado para coletar a opinião dos usuários.

HIDRATE SPARK 2.0: APLICATIVO E GARRAFA

A *Hidrate Spark 2.0* é uma garrafa de água inteligente com conexão *bluetooth* que visa ajudar o usuário a ter uma hidratação adequada, controlando, automaticamente, a sua ingestão diária de água. A garrafa também possui outros sensores que calculam a quantidade de água ingerida, são eles: microprocessador, e memória para conexão e armazenamento de dados. Esta garrafa possui capacidade de 710 mililitros (mL).

Para começar a usar a garrafa, primeiramente, é preciso fazer o download do aplicativo na *App Store* ou *Play Store*. Em seguida, deve-se criar um cadastro (perfil) no aplicativo, inserindo informações como peso, data de nascimento, nível de atividade física e propósito de uso. Tendo um perfil cadastrado, o aplicativo calcula a meta diária de ingestão de água específica para cada usuário, que pode ser personalizada a qualquer momento, e ainda o histórico de consumo de água diário e mensal.

O aplicativo em sua tela inicial (Figura 1a) possui um grande círculo azul com informações sobre a meta diária de ingestão de água (em mL) e o progresso do usuário ao longo do dia. Logo abaixo, ele mostra a porcentagem de água que o usuário alcançou durante o dia, uma estimativa de quantas garrafas ainda precisam ser bebidas e a classificação de desempenho dos últimos sete dias entre o usuário e seus amigos que possuem a garrafa. Na parte superior da tela, há informações básicas sobre a atividade do usuário ao longo da semana. A Figura 1b mostra a tela de histórico mensal da ingestão de água do usuário.

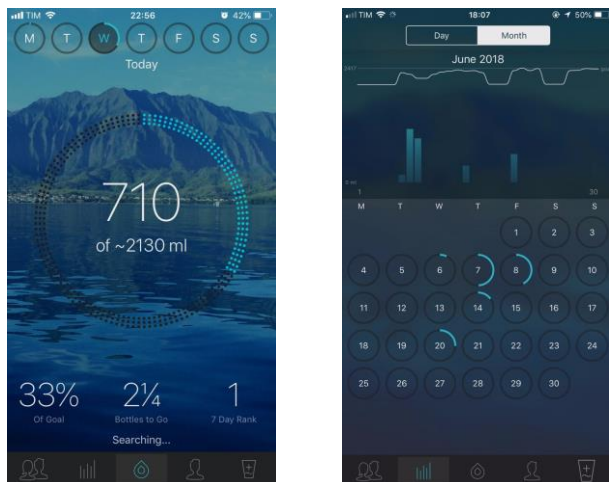


Figura 1. a) Tela inicial b) Tela de histórico mensal

MÉTODO DO ESTUDO

Para avaliar o aplicativo *Hidrate Spark*, considerando a sua comunicação com a garrafa, foram estabelecidas quatro tarefas: **T1 – Instalar o aplicativo e sincronizá-lo com a**

garrafa; T2 – Consultar o consumo de água, comparando o dia anterior e o atual; T3 – Consultar o histórico do consumo de água; e T4 – Alterar um dado do perfil do usuário (peso, altura ou nível de atividade física), a fim de verificar se a meta diária foi adaptada. Essas tarefas caracterizam as primeiras interações do usuário com o aplicativo.

Os métodos aplicados durante a avaliação se dividem em três tipos: inspeção, investigação e observação de uso [1]. Inicialmente, foi executada a Avaliação Heurística (método de inspeção) [6,10]. Em seguida, foi aplicado o MALTU (método de investigação) [5]. Por fim, foi executado o Teste de Usabilidade (método de observação de uso) [1] e uma entrevista pós-teste, a fim de coletar a opinião dos usuários sobre o uso do aplicativo e da garrafa. Os métodos utilizados são apresentados nas próximas subseções.

Avaliação Heurística

A Avaliação Heurística (AH) é um método de inspeção criado para avaliar interfaces de usuário com o objetivo de identificar problemas de usabilidade [1,6]. Nesse método,

Na Avaliação Heurística, os avaliadores inspecionam uma interface previamente selecionada a partir de um conjunto de princípios gerais de usabilidade ou heurísticas de usabilidade, e ainda classificam a severidade dos problemas encontrados, a fim de priorizar aqueles que são mais graves para a interação do usuário com o sistema. Por fim, os avaliadores se reúnem para consolidar os problemas que serão reportados no relatório de aplicação do método [6].

No presente artigo, foram utilizadas heurísticas ubíquas (HU) apropriadas para a avaliação de sistemas ubíquos [10], visto que as heurísticas iniciais propostas por Nielsen [8] podem não auxiliar na detecção de problemas relacionados com as características específicas desses sistemas, tais como: atenção, calma e transparência [3]. São elas:

Heurísticas de Usabilidade para Avaliar Sistemas Ubíquos

- **HU1** – Visibilidade do status do sistema.
Características: Calma, Atenção e Visualização das informações.
- **HU2** – Correspondência entre o sistema e o mundo real.
Características: Familiaridade, Visualização das informações, Posicionamento dos componentes, Previsibilidade, Facilidade de uso, e Usabilidade.
- **HU3** – Controle do usuário e liberdade.
Características: Calma, Aceitabilidade, Satisfação do usuário, Atenção e Reversibilidade.
- **HU4** – Consistência e Padrões.
Características: Usabilidade, Previsibilidade, Entrada de dados e Familiaridade.
- **HU5** – Prevenção de erros.
Características: Previsibilidade, Flexibilidade, Confiabilidade, Escalabilidade, Proteção e Disponibilidade.

- **HU6** – Reconhecimento ao invés de memorizar.
Características: Visualização das informações, Usabilidade, Previsibilidade, Calma e Atenção.
- **HU7** – Flexibilidade e Eficiência do uso.
Características: Flexibilidade, Utilidade, Eficiência e Eficácia.
- **HU8** – Estética e Design minimalista.
Características: Simplicidade, Calma, Atenção e Posicionamento dos componentes.
- **HU9** – Ajude os usuários a reconhecer, diagnosticar e se recuperar de erros.
Características: Simplicidade, Calma e Aceitabilidade.
- **HU10** – Ajuda e documentação.
Características: Usabilidade e Facilidade de uso.
- **HU11** – Mobilidade e dispositivos.
Características: Capacidade do dispositivo, Capacidade da rede, Interconectividade e Mobilidade.
- **HU12** – Privacidade e Segurança.
Características: Privacidade, Proteção, Segurança e Confiança.
- **HU13** – Invisibilidade e Transparência.
Características: Transparência e Previsibilidade.
- **HU14** – Sensibilidade ao contexto e interfaces adaptativas.
Características: Sensibilidade ao contexto, Adaptação e Facilidade de uso.
- **HU15** – Sensores e entrada de dados.
Características: Entrada de dados e Robustez.

MALTU

O MALTU (Modelo para Avaliação de interação em sistemas sociais a partir da Linguagem Textual do Usuário) engloba dois tipos de avaliação: usabilidade e experiência de usuário (UUX) [5]. A avaliação é feita a partir de textos espontâneos relacionados ao uso do sistema postados por usuários em Sistemas Sociais (SS). O modelo pode ser utilizado de maneira manual ou automática. Para avaliação automática existe uma ferramenta chamada de UUX-POST⁵ que classifica e analisa as Postagens Relacionadas ao Uso (PRUs) nele inseridas.

A metodologia de avaliação do MALTU tem como objetivo guiar um profissional de Interação Humano-Computador (IHC) na avaliação de um SS a partir de um conjunto de comentários (PRUs). A metodologia explica as cinco atividades de um avaliador, representadas em etapas sequenciadas para avaliação.

A primeira etapa, definição do contexto de avaliação, é subdividida em três partes: 1) contexto de uso do sistema; 2) domínio do sistema; 3) objetivos de avaliação. O contexto de uso do sistema segundo o MALTU consiste em definir plataforma e usuários. Neste trabalho, o usuário foi definido de maneira geral como sendo: usuários de dispositivos ubíquos. A maioria dos dispositivos ubíquos utiliza a plataforma mobile para comunicação, dessa forma,

fica estabelecido que o dispositivo a ser avaliado deve oferecer suporte a pelo menos dois sistemas operacionais mobile mais populares: Android e iOS. A segunda atividade, denominada de extração das PRUs, pode ser feita de duas formas: extração manual ou automática. Neste trabalho, utilizamos a maneira manual, a qual consiste na coleta e armazenamento em formato CSV (Excel) das PRUs do fórum de discussão do dispositivo. A terceira etapa envolve a classificação das PRUs, que foi realizada de maneira manual por dois especialistas em IHC, os quais analisaram e classificaram as PRUs em tipos como elogio, crítica, dúvida, comparação e sugestão. Na quarta etapa, os resultados devem ser organizados e exibidos obedecendo uma relação lógica entre eles. Deve-se gerar gráficos e tabelas com exemplos, objetivando o melhor entendimento dos interessados na avaliação. Na última etapa, deve-se relatar os resultados.

Teste de Usabilidade

O Teste de Usabilidade é um método de observação de uso, cujo objetivo é avaliar a interação a partir da experiência de uso dos usuários com o sistema interativo. Este método é aplicado através do monitoramento das ações do usuário no sistema escolhido por meio de atividades previamente definidas. Estas tarefas possuem a finalidade de testar as funcionalidades-chave do sistema a partir do ponto de vista do usuário, com isso as tarefas agem como condutoras do teste [1].

A aplicação deste método segue algumas etapas [1]: a primeira é a definição de quem são os usuários-alvo. Em seguida, deve-se decidir se o teste será executado em laboratório ou no ambiente real do usuário [11]. Posteriormente, é necessário definir as tarefas e métricas que serão abordadas e avaliadas, considerando o objetivo da avaliação. Por fim, recomenda-se a aplicação de um teste piloto, a fim de calcular o tempo médio para a execução de cada tarefa e para validar os materiais elaborados para a execução do teste, incluindo as perguntas elaboradas para a execução de uma entrevista pós-teste.

No presente artigo, o objetivo estabelecido para o Teste de Usabilidade foi avaliar a facilidade de aprendizado e a experiência de uso nas primeiras interações com o aplicativo *Hidrate Spark* e sua garrafa. Os usuários-alvo são pessoas com idade acima de 16 anos, que se preocupam em manter hábitos saudáveis, especialmente em relação à ingestão de água, e que possuem conhecimento básico sobre a língua inglesa, visto que é o único idioma disponível no aplicativo.

Para a aplicação do teste foram estabelecidas quatro tarefas a serem executadas pelos usuários em ambiente controlado. São elas: TU1 – Instale o aplicativo, faça o cadastro de usuário e sincronize o dispositivo com a garrafa; TU2 – Informe qual é a sua meta diária de ingestão de água calculada pelo aplicativo; TU3 – Simule a ingestão de água e informe quanto falta para concluir a meta diária; e por fim, TU4 – Acesse as configurações de perfil do usuário e

⁵ <http://uuxposts.russas.ufc.br>

altere um dado importante para o cálculo da meta diária (peso, altura ou nível de atividade física), e após isso, informe qual é a sua nova meta diária. Estas tarefas estão relacionadas com as tarefas planejadas para toda a avaliação (apresentadas anteriormente). As tarefas TU2 e TU3 não possuem a mesma descrição da tarefa T2, porém elas estão associadas com a funcionalidade de consultar a meta diária de consumo de água. A tarefa T3 é a única que não foi solicitada aos usuários, visto que, para a execução dessa atividade, os usuários precisariam usar o aplicativo e a garrafa por pelo menos dois dias.

As métricas consideradas para a execução do teste são: (i) quantidade de tempo para realizar a tarefa; (ii) número de vezes que o usuário solicitou ajuda; e (iii) quantidade de erros. Para auxiliar na análise das métricas o áudio de cada participante do teste foi gravado.

O roteiro do teste foi planejado e executado da seguinte maneira: inicialmente, os pesquisadores se apresentaram para os participantes. Depois, eles explicaram a proposta da garrafa e do aplicativo *Hidrate Spark*, e o objetivo do teste. Em seguida, foi feita a leitura do Termo de Consentimento junto com os participantes, e eles assinaram o termo. Após isso, o teste foi aplicado, e por último, foi realizada uma entrevista pós-teste.

A entrevista é um método de investigação bastante utilizado para conhecer fatos sobre as pessoas, permitindo a coleta de dados para análises científicas [4]. Foi aplicado após o Teste de Usabilidade com o intuito de obter a opinião dos usuários após o uso.

APLICAÇÃO DA AVALIAÇÃO HEURÍSTICA

A Avaliação Heurística foi executada considerando as tarefas definidas na seção MÉTODO DO ESTUDO e envolveu a participação de cinco avaliadores, que utilizaram *smartphones* com o sistema operacional (SO) Android e iOS. O nível de experiência dos avaliadores com este método e os dispositivos utilizados por cada um são apresentados na Tabela 1. O avaliador A1 executou a inspeção em 31 dias, o A2 em 6 dias, o A3 em 4 dias e os avaliadores A4 e A5, cada um, em 2 dias.

Avaliador	Experiência	Dispositivo	SO
A1	Experiente	Samsung Galaxy S8	Android 6.0.1
A2	Experiente	Motorola Moto Z ² Play	Android 8.0.0
A3	Iniciante	Apple iPhone 6S	iOS 12.0 Beta
A4	Iniciante	Apple iPhone 7 Plus	iOS 11.4
A5	Iniciante	Motorola Moto G 2 ^o Geração	Android 7.1.2

Tabela 1. Experiência dos avaliadores e dispositivos utilizados

Os avaliadores identificaram um total de 33 problemas de usabilidade na interface do aplicativo, em que 17 deles foram encontrados somente no Android, 3 no iOS e os outros 13 em ambos os sistemas operacionais, como mostra a Tabela 2. É bem verdade que os avaliadores mais experientes inspecionaram o aplicativo Android. Entretanto, durante a consolidação dos problemas, esse fato foi discutido e ficou claro que a experiência de uso fornecida pelas duas plataformas (Android e iOS) é bastante distinta.

Tarefa	Nº de problemas identificados	Android	iOS	Ambos
T1	11	11	0	0
T2	11	2	1	8
T3	8	3	1	4
T4	3	1	1	1
Total	33	17	3	13

Tabela 2. Número de problemas identificados para cada tarefa e sistema operacional

As tarefas T1 e T2 apresentaram uma maior ocorrência de problemas (11 para cada uma). Os problemas da tarefa T1, identificados apenas no Android, afetam as telas de cadastro do usuário, como é o caso em que o aplicativo não mantém a data de nascimento quando o usuário retorna de telas posteriores para a tela de seleção da data de nascimento, apresentando um valor padrão. E os problemas referentes à tarefa T2 prejudicam o uso da tela inicial, que mostra o consumo diário de água do usuário, como é o caso em que o usuário ficou em dúvida se o aplicativo registrou o seu consumo de água quando a garrafa não estava posicionada em uma superfície plana, já que ela foi usada em um veículo em movimento.

A tarefa T3 apresentou 8 problemas, que afetam as telas de histórico do consumo de água diário e mensal. Por exemplo, o gráfico do consumo de água diário mostrou que o usuário consumiu água em um horário que ele não ingeriu água, de acordo com os horários apresentados nos registros abaixo do gráfico. Por fim, a tarefa T4, que está relacionada com a tela de informações pessoais do usuário e com a tela inicial, apresentou apenas 3 problemas, o menor número registrado durante a inspeção, como é o caso em que um dado importante para o cálculo da meta diária de ingestão de água (peso, altura ou nível de atividade física) foi editado, porém a meta só foi alterada ao reiniciar o aplicativo.

Baseado na escala de severidade de Nielsen [7], foram identificados 6 problemas cosméticos (4 no Android e 2 em ambas as plataformas), como é o caso em que o teclado do Android é escondido automaticamente quando o usuário está digitando seu nome e decide apagar todo o texto clicando continuamente no botão de *backspace*. Em relação aos problemas pequenos, foram identificados 11 problemas (5 no Android, 1 somente no iOS e 5 nas duas plataformas).

Por exemplo, o caso em que o usuário está inserindo o e-mail e o teclado esconde o campo de senha, deixando o usuário perdido sem saber onde inserir essa informação.

Sobre os problemas grandes, foram encontrados 15 problemas (7 no Android, 2 somente no iOS e 6 em ambas as plataformas), como é o caso em que o aplicativo sincroniza o consumo de água do usuário, mostra a informação rapidamente e depois retira ela da tela. Em relação aos problemas catastróficos, identificou-se apenas um problema no Android, que o ocorre na tela de finalização do cadastro, em que o aplicativo não fornece uma opção de “sign in” ou “esqueci minha senha” quando o usuário já possui um perfil cadastrado.

As heurísticas de usabilidade mais violadas foram: a HU1 (Visibilidade do status do sistema) com 11 problemas, em que 6 deles ocorreram somente no Android, 1 apenas no iOS e 4 em ambas as plataformas; a HU7 (Flexibilidade e Eficiência do uso) com 10 problemas, em que 7 deles apareceram apenas no Android, 2 somente no iOS e 1 nas duas plataformas; e a HU4 (Consistência e Padrões) com 9 problemas, em que 6 deles ocorreram somente no Android e os outros 3 nas duas plataformas. Isso comprova o fato de que o aplicativo Android possui menor qualidade de uso, visto que foram identificados mais problemas nesta plataforma em comparação com o iOS. Essa análise também indica que a interface dos aplicativos não mantém o estado do sistema e não fornece feedback às ações dos usuários de forma apropriada; a completude das tarefas é prejudicada por não haver boa eficiência de uso; as informações sofrem inconsistência; e não são usados padrões familiares ao usuário, que possibilitem um melhor entendimento, aprendizado e uso da interface.

Considerando apenas as heurísticas mais relacionadas com as características específicas dos sistemas ubíquos (HU11-HU15), as mais violadas foram: a HU11 (Mobilidade e dispositivos) e a HU14 (Sensibilidade ao contexto e interfaces adaptativas). Os aplicativos das duas plataformas apresentam defeitos em relação às características ubíquas. Essa análise também mostra que os aplicativos não mantêm algumas funcionalidades funcionando corretamente quando ocorre deslocamento físico dos usuários e que existem problemas relacionados com as propriedades dos dispositivos, como a comunicação entre o aplicativo e a garrafa através do *bluetooth*, a capacidade de bateria do *smartphone* e a capacidade de memória da garrafa. Por fim, os aplicativos não adaptam o uso da interface a partir da mudança de contexto dos usuários e dos dispositivos.

As características ubíquas mais afetadas foram: Atenção, Interconectividade e Sensibilidade ao contexto (5 problemas para cada uma). Isso comprova a existência dos problemas relacionados com a transmissão de informações por meio da comunicação entre os dispositivos, e com a adaptação das funcionalidades a partir das mudanças contextuais. Ressalta-se também que o usuário pode esquecer sua atividade principal (ingestão adequada de

água) pelo fato de se preocupar com a tecnologia envolvida no sistema ubíquo (característica de Atenção).

A listagem completa dos problemas identificados através da Avaliação Heurística e as telas analisadas podem ser acessadas *online*⁶.

APLICAÇÃO DO MALTU

Através do modelo MALTU foram analisadas 99 PRUs (21 foram descartadas por não terem informações relevantes à análise), retiradas da plataforma Amazon.com, onde além de ser vendida a garrafa *Hidrate Spark 2.0*, há um local destinado aos compradores da garrafa, para eles opinarem sobre ela baseado na sua experiência. Além disso, existem comentários de usuários do iOS e Android. O critério para extração das PRUs foram comentários que possuíam até três estrelas atribuídas, sendo divididas em três partes iguais. Nesses critérios, ficaram 33 PRUs com uma estrela, 33 PRUs com duas estrelas e 33 PRUs com três estrelas. Inicialmente esses comentários estavam em Inglês, mas sofreram uma tradução automática pelo Google Tradutor. Quando os avaliadores tiveram dúvidas quanto à tradução, a versão original foi consultada. Partindo da extração para a classificação, as PRUs eram categorizadas em tipo (crítica, elogio, dúvida, comparação, sugestão e ajuda), intenção (visceral, comportamental e reflexiva), usabilidade (eficiência, satisfação, segurança, utilidade, memorabilidade e aprendizado) e experiência de usuário (satisfação, afeto, confiança, estética, frustração, motivação e suporte).

	Categoria	Nº de PRUs
Tipo	Crítica	76 (97.5%)
	Sugestão	2 (2.5%)
Intenção	Visceral	54 (69.2%)
	Comportamental	12 (15.4%)
	Reflexiva	12 (15.4%)
Funcionalidade	Bateria	35 (44.8%)
	Função principal	23 (29.5%)
	Conexão	17 (21.9%)
	Luz da garrafa	3 (3.8%)
Usabilidade	Satisfação	40 (51.7%)
	Eficiência	24 (30.6%)
	Eficácia	8 (10.1%)
	Utilidade	4 (5%)
	Aprendizado	2 (2.6%)
UX	Frustração	55 (70.5%)
	Satisfação	22 (28.2%)
	Confiança	1 (1.3%)

Tabela 3. Classificação das PRUs através das categorias

⁶ <http://bit.ly/2CGyFzr>

Com os resultados obtidos é possível observar que a maioria das PRUs válidas são críticas relacionadas com a função principal (registro do consumo de água), conexão da garrafa com o aplicativo, bateria e a luz emitida pela garrafa. Somente 2 usuários sugeriram algo. As PRUs também receberam classificação quanto as intenções dos usuários, essa categorização mede o quanto o usuário apresenta detalhes da funcionalidade que ele está criticando e a intensidade de sentimento. A visceral foi a que mais apareceu na classificação em aproximadamente 70% das PRUs (Tabela 3). Essa classificação é definida com alta intensidade de sentimento e pouco detalhamento do problema. A comportamental contém média intensidade de sentimento e detalhamento representou 14.2% das PRUs. Já a reflexiva é definida com baixa intensidade de sentimento e alto detalhamento e ficou com a mesma porcentagem da comportamental. Na categoria funcionalidades, o maior problema encontrado foi com relação a bateria (44.8%) seguido da principal função do sistema, que é o registro do consumo de água (29.5%). Outro problema bem comum relatado foi a conexão com o aplicativo: 21.9% dos comentários relataram problemas com o emparelhamento via *bluetooth*. Com relação à usabilidade, destacaram-se negativamente a satisfação, eficiência, eficácia, utilidade e aprendizado do sistema. Em satisfação, onde é relatado o grau de aceitação do usuário e quão o sistema é agradável a ele, 51.7% das PRUs relatavam esse problema, 30.6% criticaram a eficiência, relatando um nível de dificuldade elevado para realizar as tarefas. Apenas 10.1% dos usuários não conseguiram usufruir da garrafa por não conseguirem utilizar a função principal (problema de eficácia). A garrafa e o aplicativo de maneira geral possuem uma boa aprendizagem, sendo que somente 2.6% das PRUs relatavam problemas dessa categoria. No quesito UX, apenas três categorias foram identificadas. São elas: frustração, satisfação e confiança. Grande parte dos usuários (70.5%) se sentiram frustrados com o sistema, 28.2% das PRUs relatavam problemas de satisfação e uma única PRU relatou problema de confiança na utilização da garrafa.

APLICAÇÃO DO TESTE DE USABILIDADE

A execução do Teste de Usabilidade ocorreu em ambiente controlado. Inicialmente, realizou-se um teste piloto, em que se identificou a necessidade de disponibilizar aos usuários uma garrafa simples contendo água para a realização da tarefa de simulação de ingestão de água (TU3). Em seguida, seis usuários foram recrutados considerando o perfil estabelecido (Tabela 4). Sobre os dispositivos utilizados, quatro participantes usaram *smartphone* iOS e os outros dois utilizaram Android.

Usuário	Sexo	Faixa Etária	SO
U1	Masculino	18-22	iOS
U2	Masculino	18-22	iOS
U3	Feminino	23-27	Android

U4	Feminino	18-22	Android
U5	Masculino	23-27	iOS
U6	Feminino	23-27	iOS

Tabela 4. Dados dos usuários participantes do teste

Para apoiar a análise das métricas coletadas foram estimados valores referentes à execução das tarefas. Esses números foram baseados na experiência dos avaliadores com o aplicativo através da execução da Avaliação Heurística. O tempo estimado é o mesmo para as plataformas Android e iOS, com exceção da tarefa TU4, em que foram calculados tempos diferentes para cada sistema operacional, visto que a execução dessa tarefa envolve telas diferentes dependendo da plataforma.

O tempo estimado da tarefa TU1 (Instalação do aplicativo, cadastro e sincronização com a garrafa) foi 157 segundos. Os usuários gastaram mais tempo que o esperado para executar essa tarefa. Isso também foi observado na tarefa TU2 (Consulta da meta diária de ingestão de água), que teve tempo estimado de 15 segundos. O usuário U2 foi o único que se aproximou da estimativa, gastando 17 segundos, e a usuária U4 foi quem mais se distanciou do esperado, gastando 231 segundos.

O tempo estimado da tarefa TU3 (Simulação da ingestão de água e consulta da nova meta diária) foi 76 segundos. O usuário U4 realizou a tarefa com tempo aproximado de 88 segundos. Os usuários U1 e U2 gastaram menos tempo que o esperado (71 e 44 segundos respectivamente). Entretanto, os outros usuários não conseguiram executar a tarefa no tempo esperado, variando entre 100 e 172 segundos. O tempo médio de execução dessa tarefa foi 102.33 segundos.

O tempo estimado da tarefa TU4 (Alteração de um dado do perfil do usuário para modificar a meta diária) foi 57 segundos na plataforma iOS e 44 segundos na plataforma Android. Um dos usuários do iOS (U1) se aproximou da estimativa, gastando 61 segundos. Os outros usuários do iOS realizaram a tarefa entre 128 e 250 segundos, indicando que eles gastaram mais tempo que o esperado. Um dos usuários do Android (U3) também se aproximou desse valor, gastando 219 segundos. Contudo, a outra usuária (U4) do Android se aproximou do esperado, gastando 47 segundos. O tempo médio de execução dessa tarefa foi 147 segundos.

A situação estimada seria que os usuários não solicitassem ajuda e que não ocorressem erros na execução das tarefas. Entretanto, na Tabela 5 pode ser visto que foram feitas 53 solicitações de ajuda.

Usuários / Tarefas	TU1	TU2	TU3	TU4	Total
U1	4	0	2	0	6
U2	4	3	2	1	10
U3	4	0	2	1	7

U4	4	2	0	0	6
U5	7	2	4	0	13
U6	4	3	2	2	11
Total	27	10	12	4	53

Tabela 5. Número de solicitações de ajuda durante a execução das tarefas

Os participantes que usaram dispositivo Android requisitaram ajuda 6 e 7 vezes respectivamente, enquanto que alguns usuários do iOS solicitaram ajuda entre 10 e 13 vezes. Apenas um usuário do iOS (U1) requisitou ajuda 6 vezes. Isso pode indicar que os usuários do Android executaram as tarefas com maior facilidade que alguns usuários do iOS, contudo o número de usuários das duas plataformas não foi balanceado, impedindo uma melhor comparação. Os usuários que possuem entre 18 e 22 anos solicitaram ajuda entre 6 e 10 vezes, enquanto que os usuários com idade entre 23 e 27 anos requisitaram ajuda entre 7 e 13 vezes, indicando que essa variável também não influenciou nas dificuldades dos usuários. Em geral, essa análise reforça o fato de que os participantes tiveram dúvidas e dificuldades ao usar a interface do aplicativo.

Em relação a quantidade de erros ocorridos na execução das tarefas, ocorreram 19 erros na tarefa TU1, como mostra a Tabela 6.

Usuários / Tarefas	TU1	TU2	TU3	TU4	Total
U1	3	0	0	3	6
U2	3	0	1	2	6
U3	3	0	1	2	6
U4	3	0	0	0	3
U5	5	3	3	5	16
U6	2	0	2	0	4
Total	19	3	7	12	41

Tabela 6. Quantidade de erros durante a execução das tarefas

Considerando os participantes que usaram dispositivo iOS, o número de erros foi 6 para os usuários U1 e U2, 4 para a usuária U6 e 16 para o usuário U5 (caso discrepante). Para os usuários do Android (U4 e U3) foram identificados 3 e 6 erros respectivamente. Na análise dessa métrica percebe-se que o sistema operacional utilizado não influenciou na realização das tarefas. Para cada usuário homem foram identificados 6 erros, exceto o usuário U5 que é o caso discrepante, e para as mulheres o número de erros variou entre 3 e 6, mostrando que essa variável não influenciou na quantidade de erros ocorridos. Para os usuários que possuem de 18 a 22 anos o número de erros variou entre 3 e 6 e para os usuários (U4 e U6) com idade entre 23 e 27 anos o número de erros foi 3 e 4 respectivamente, excluindo o usuário U5 (discrepante). Desse modo, a faixa etária também não influenciou na quantidade de erros ocorridos.

Em geral, essa análise indica a existência de interações erradas durante a execução das tarefas.

A análise das métricas ainda indica que os valores ideais definidos não foram alcançados pela maioria dos usuários, reforçando a dificuldade de uso da interface. A seguir são apresentados alguns problemas observados pelos aplicadores do teste de usabilidade.

Na tarefa TU1 (Instalação do aplicativo, cadastro e sincronização com a garrafa), os usuários tiveram dificuldade em entender quais dados o aplicativo estava solicitando, como mostram alguns comentários dos usuários: “O nome é completo ou é...” e “Hmm... entendi não” do usuário U1; “Eu faço o que?” e “E agora?” da usuária U3; “Aqui eu coloco meu e-mail ou o nome mesmo?” do usuário U5; por fim, “Bota o nome completo ou só o primeiro nome?” da usuária U6.

Os usuários tiveram dificuldade em compreender as instruções de uso da garrafa dadas pelo aplicativo, muitas vezes eles acharam que precisavam realizar alguma interação. Isso pode ser observado através dos seguintes comentários: “E agora?” do usuário U1; “É pra encher? Eu não sei, eu tô com medo até de fazer alguma coisa aqui” do usuário U2, representando insegurança no uso; “É pra reiniciar é?” da usuária U3; “Essa parte aqui...” e “Tem que fazer alguma coisa?” da usuária U4; “Aqui tá perguntando se?” e “Isso aqui é instrução se for encher a garrafa? É ou não?” do usuário U5; “Me explica isso aqui”, “Aqui tá mandando eu encher, é isso?” e “São só instruções aqui, é isso?” da usuária U6. A dificuldade dos usuários em entender as informações do aplicativo pode ter ocorrido devido ao idioma (Inglês). Por exemplo, o usuário U5 disse: “Olha as perguntas em inglês para quem não sabe inglês”.

Os usuários não entenderam o comportamento do aplicativo quando ele estava sincronizando o dispositivo com a garrafa, como mostram os seguintes comentários: “Aqui agora é o que?” do usuário U1 e “O que tá acontecendo aqui?” do usuário U5. Isso fez com que a usuária U5 tivesse dúvida sobre o *bluetooth*, que precisava estar ligado para a conexão entre os dispositivos, ela disse: “Pera aí, tem que ligar o *bluetooth* né?”. Alguns usuários tiveram dificuldade para sincronizar o dispositivo com a garrafa, como mostra os comentários: “Dá uma olhada aqui pra ver se conectou, se não tá aparecendo...” da usuária U3 e “Tá tipo processando aqui!” do usuário do usuário U5. A usuária U3 conseguiu conectar os dispositivos, gastando cerca de 4 minutos e 48 segundos apenas para realizar essa atividade, porém o usuário U5 não conseguiu conectar de nenhuma maneira, gastando cerca de 2 minutos e 43 segundos nessa atividade, por isso, ele passou a usar um dispositivo iOS e reiniciou a tarefa.

Outros problemas referentes à tarefa TU1 também foram identificados. A usuária U3 não entendeu que a localização influencia no cálculo da meta diária de ingestão de água, por isso, ela ficou em dúvida se deveria permitir o uso do

sensor de localização, ela disse: “Eu permito a localização do dispositivo? Ele [aplicativo] tá perguntando”. A usuária U6 não sabia que poderia diminuir a altura arrastando a barra lateral localizada na tela de escolha da altura, ela disse: “Como é que eu faço para diminuir minha altura?”. O aplicativo mostra um cronômetro estático e pede que o usuário espere 10 segundos para medir corretamente a quantidade de água existente na garrafa. Contudo, o usuário U5 entende que o cronômetro é dinâmico e que vai contabilizar o tempo. Os condutores do teste perceberam o problema e perguntaram ao usuário se ocorreu algo de errado, ele disse: “Não, estou só esperando o tempo, né!”, “Não sei se já passou de 10 segundos” e “Não passou nem um”. Ressalta-se que o usuário U2 começou a tarefa com um dispositivo Android, que travou e foi trocado por um iOS; assim, o usuário reiniciou a tarefa. As dificuldades dos usuários justificam o fato de a tarefa TU1 possuir o maior tempo médio (450.33 segundos), maior número de solicitações de ajuda (4 vezes) e maior quantidade de erros (19). Ressalta-se que apenas os dispositivos Android usados pelas usuárias U3 e U4 não estavam com o *bluetooth* ativado, por isso, esse problema não foi relatado pelos usuários do iOS.

Na tarefa TU2, os usuários tiveram dificuldade em localizar a meta diária de ingestão de água em mL, como mostram os comentários: “Seria aqui?” do usuário U1; “Essa daqui. Não é essa não?” da usuária U4; e, “É isso aqui, ou não?” do usuário U5. O usuário U1 pensou que a meta era o número de garrafas a serem bebidas por dia, ele perguntou: “*Bottles to Go* quer dizer o que?”, indicando dificuldade com o idioma. Após saber o significado, ele disse: “Deve ser isso aqui né a meta?”. Na terceira tentativa ele encontrou e disse: “Eu fiquei perdido”. Esse problema também ocorreu com a usuária U4, em que um condutor perguntou se ela achava que a meta deveria ser informada em mL ou quantidade de garrafas, ela disse: “Em mL? É melhor em mL. Só que mL...”, mostrando que ela não tinha encontrado a meta. Em outra tentativa ela disse: “Eu achei aqui”. O usuário U5 encontrou com mais facilidade, mas teve dúvida e disse: “A meta não é isso aqui que eu atingi não né? De dois litros e alguma coisa?”. Após a confirmação do condutor, ele disse: “Ah certo, eu procurando, achei que tinha que calcular algum número. [...], mas ele já dá a meta aqui né”. Os usuários U1, U4 e U5 podem ter gastado mais tempo nessa tarefa por causa das dificuldades enfrentadas. Além de que o usuário U5 foi o único com erros na tarefa.

Na tarefa TU3 (Simulação da ingestão de água e consulta da nova meta diária), alguns usuários não entenderam que o aplicativo calcula a quantidade de água restante para atingir a meta diária, como mostram os comentários: “Em números ou em..., quer dizer que eu tenho que fazer o cálculo?” (U1), e “Em números? É, eu acho que... não tô achando direito em números, mas fazendo a subtração” (U4). O usuário U5 compreendeu melhor as instruções da tarefa TU1: “Ah certo, as instruções são para fazer isso todas as

vezes né”. Ele ficou em dúvida se precisava fazer algo a mais para o aplicativo registrar o consumo; ele disse: “Precisa fazer alguma coisa aqui ou não?”, isso mostra que a conexão dos dispositivos precisa ser mais rápida. O aplicativo registrou 220 mL para a usuária U6 antes de a simulação começar. A dificuldade enfrentada pelo usuário U5 pode ser o motivo para ele ter o maior número de solicitações de ajuda (4 vezes) e a maior quantidade de erros (3), além de que ele foi um dos participantes que gastou mais tempo (100 segundos) nessa tarefa.

Na tarefa TU4, os usuários editaram um dado do perfil do usuário (peso, altura ou nível de atividade física), porém a meta diária não foi atualizada, como mostram os comentários: “E agora?” (U1); “Alterei, mas ele não mudou lá não, mas eu acredito que eu tenha alterado. Mas pelo o que vi aqui, não apareceu nenhuma diferença. Tá atualizando? Eu acho que não” (U2); por fim, “Vou tentar de novo. Não, ainda continua a mesma” (U6). Os usuários U1 e U6 tiveram a meta modificada após reiniciar o aplicativo, a usuária U6 disse: “Pronto, agora alterou”. O usuário U2 não completou a tarefa. Isso pode ter ocorrido porque ele não reiniciou o aplicativo. Esse problema ocorreu apenas para os usuários do iOS, comprovando que eles tiveram mais dificuldade que os usuários do Android nessa tarefa. Em relação à alteração do dado, o usuário U5 clicou sobre o nome “Peso” e não em cima do valor correspondente, assim, o aplicativo não fez nada. No aplicativo Android, existe uma aba sobre o perfil do usuário e outra apenas para configurações. Isso gerou dúvida para a usuária U3 que disse: “Onde é que eu clico aqui para ir para as configurações do meu perfil? É aqui mesmo?”, referenciando a aba de perfil e não a de configurações. Os usuários U2 e U6 estão entre os participantes que gastaram mais tempo (177 e 250 segundos respectivamente) na tarefa, e o U2 foi quem mais solicitou ajuda (2 vezes). Isso pode ter acontecido por causa das dificuldades enfrentadas por eles na tarefa. O problema da usuária U3 também pode justificar o longo tempo gasto na tarefa (219 segundos).

Aplicação da Entrevista Pós-teste

Uma entrevista pós-teste foi aplicada para coletar a opinião dos usuários. Eles gostaram da ideia e disseram que o uso da garrafa e do aplicativo é prático, como mostram as seguintes respostas: “Eu acho bem prático, a ideia eu acho bem interessante” (U1); “Bonito o design. Conexão com ela é bem rápido também” (U2); “Eu achei interessante, eu gostei, eu gostei” (U3); “Achei bem criativa e útil também” (U5); “Eu gostei, quero uma garrafa dessa agora. Eu acho que ela é bem útil” (U6).

Somente a usuária U4 não achou a ideia útil, ela disse: “Eu acho que é uma ideia boa, mas talvez não seria tão útil. Eu acho que é porque não tem muita praticidade, não sei, tem que tá conectado... Eu acho que interessa mais pra quem realmente quer cuidar da saúde”. Contudo, acredita-se que houve mudança de ideia, visto que ela falou que usaria a garrafa e justificou dizendo: “Pra mim é muito pessoal,

porque eu vejo que eu não tomo muita água, eu acho que seria bem útil”. Os outros usuários também usariam a garrafa, principalmente, para aprimorar o cuidado com a saúde, por exemplo:

- “Usaria, porque eu acho que influencia a gente a fazer algo que a gente precisa fazer e não faz normalmente” (Usuário U1);
- “Sim. Me incentivar no consumo de água, que eu bebo pouca água. É por isso que ela é interessante também, porque vai te ajudar a controlar como você tá no dia-a-dia, como é que você tá bebendo água, se não tá. É por isso que eu achei interessante” (Usuária U3);
- “Se eu tivesse uma garrafinha dessa eu usaria com certeza. Bom, é porque eu acabo esquecendo de ingerir a quantidade correta de água, tipo lá na sala todo mundo tem a garrafinha, nem a garrafinha eu tenho” (Usuário U5).

Apenas o usuário U2 disse que não usaria a garrafa, ele justificou dizendo: “Se fosse um pouco menor, acho que um pouco menor eu usaria. Esse tamanho não”.

Os usuários disseram que entenderam como a garrafa e o aplicativo funcionam. Contudo, alguns não compreenderam, inicialmente, o funcionamento com clareza, como pode ser observado nas seguintes respostas: “Entendi... entendi por cima” (U1); “Sim, sim, entendi. A primeira parte eu não tinha entendido bem, que ela calcula a partir do momento que você tira né... água” (U3).

Sabendo-se que o aplicativo só possui a língua inglesa como idioma, os condutores tentaram evitar o problema recrutando somente participantes que afirmaram conhecer o básico do idioma. Contudo, nota-se que os participantes precisariam ter um domínio maior da língua, por isso, os resultados do Teste de Usabilidade podem ser diferentes se aplicados com pessoas que possuem conhecimentos médios ou avançados em Inglês.

Por fim, o usuário U2 relatou outras dificuldades: “Não deixar digitar o nome, eu começava a digitar, apertava espaço, ele já fechava e não ficava com o nome que você já preencheu. Foi só voltar e abrir de novo que funcionou” (esse problema aconteceu quando ele usou o Android). Ele continuou: “E configuração eu não achei, mas eu acho que é por causa do celular porque o pontinho do iPhone ficou em cima do configuraçãozinha, aí não vi”. Esse último problema possui relação com o uso do dispositivo iOS e não com o uso do aplicativo.

CONSIDERAÇÕES FINAIS

Avaliamos o sistema *Hidrate Spark* utilizando a Avaliação Heurística, o Modelo MALTU e o Teste de Usabilidade, seguido de entrevista. Podemos dizer que todos eles foram importantes para a avaliação. Vimos que configurar o aplicativo e sincronizá-lo com a garrafa (T1) não é uma tarefa fácil, isso foi apontado tanto na Avaliação Heurística (11 problemas), como no Teste de Usabilidade e entrevista. O MALTU também apontou várias críticas sobre

a dificuldade de conexão da garrafa com o aplicativo. Quanto à consulta do consumo de água (T2), encontramos comentários no MALTU apontando dúvida sobre o registro de consumo, por exemplo: “Eles falharam em nos dizer que depois de beber água, você tem que colocar em uma superfície plana por 10 segundos para registrar a ingestão de água, o que faz de forma imprecisa.” A Avaliação Heurística nos mostrou 11 problemas relacionados a essa tarefa. O Teste de Usabilidade também confirmou a dificuldade de entender o quanto foi consumido de água no dia. Quanto à consulta do histórico do consumo de água (T3), a Avaliação Heurística revelou quatro graves problemas de usabilidade. O MALTU não apontou explicitamente nenhum dessa tarefa. O Teste de Usabilidade não conseguiu avaliar essa tarefa porque seria necessário o registro do consumo de alguns dias para que o histórico fosse acessado. Quanto a alterar um dado do perfil do usuário, enquanto a Avaliação Heurística apontou apenas 3 problemas, que foram reforçados no Teste de Usabilidade. O MALTU não registrou nenhum problema dessa natureza.

Observamos que problemas relatados no MALTU foram também identificados na Avaliação Heurística, como por exemplo a dificuldade de entender o registro do consumo de água. Problemas encontrados no Teste de Usabilidade também foram percebidos na Avaliação Heurística, a exemplo da dificuldade em entender como os dados solicitados são usados para o cálculo da meta diária de ingestão de água. Alguns problemas foram identificados apenas no Teste de Usabilidade, tais como: dificuldade para entender quais dados estão sendo solicitados e para compreender as instruções de uso da garrafa; dificuldade para entender o comportamento do aplicativo no momento de sincronização com a garrafa; o aplicativo pede que o usuário espere 10 segundos com a garrafa em superfície plana para medir corretamente que a garrafa está vazia, para isso, ele mostra um cronômetro que parece ser dinâmico e regressivo, mas é estático na verdade; dificuldade em localizar a meta diária de ingestão de água em mL; o usuário ficou em dúvida se conseguiria acessar as configurações do perfil do usuário por meio da aba de perfil ou da de configurações (ocorre apenas no Android). Vale ressaltar que o smartphone Android usado por um usuário travou completamente na primeira tentativa do Teste de Usabilidade. Esse problema não ocorreu com os avaliadores que usaram Android na Avaliação Heurística.

Outros problemas foram encontrados somente na Avaliação Heurística, tais como: o aplicativo perde as informações inseridas pelo usuário quando ele avança para uma tela e depois volta para a tela já concluída; os menus para escolha das medidas (e.g., cm ou ft) aparecem em cinza, sugerindo que não são clicáveis; ao digitar o e-mail, o teclado esconde o campo de senha, deixando o usuário perdido sem saber onde inserir esse dado; quando o usuário já possui cadastro, o aplicativo não fornece uma interação alternativa; o aplicativo não armazena corretamente o consumo de água

do usuário, às vezes perdendo alguns valores; dificuldade para entender se o aplicativo registra o consumo de água quando a garrafa está sendo usada em um veículo em movimento, quando o *bluetooth* está desligado e quando o usuário está dormindo; não existe uma opção para interromper o funcionamento da garrafa por um determinado período; a meta diária sofreu alteração de um dia para o outro sem interferência do usuário; os botões de alterar e editar um registro de consumo de água localizados na tela de histórico do consumo de água não parecem ser clicáveis, e as informações dessa tela não estão padronizadas em um único idioma; o gráfico diário de consumo de água mostra que o usuário bebeu água em um horário que ele não fez isso; e o usuário fica em dúvida se pode visualizar o gráfico diário referente aos dias passados e o gráfico mensal de consumo de água.

Observamos que, para avaliar a usabilidade em sistemas ubíquos, precisamos de “tempo” para que os problemas se revelem. Por exemplo, o problema em que o aplicativo não registrou o consumo de água do usuário, no caso em que o seu *smartphone* descarregou, só foi identificado através dessa mudança de contexto. Reforçando essa ideia, observamos que o avaliador que ficou mais tempo com a garrafa durante a Avaliação Heurística identificou problemas que os outros avaliadores não identificaram, como por exemplo, “O usuário não sabe se o aplicativo registrou o consumo de água quando a garrafa está sendo usada em um veículo em movimento”. Portanto, ficou o sentimento de que um Teste de Usabilidade para avaliar sistemas ubíquos deveria ocorrer em várias seções de observação, e mesmo assim pode ainda não revelar alguns problemas pelo fato de ser um teste de laboratório e com isso a mudança de contexto e a mobilidade poderiam não ser observadas corretamente.

Já o MALTU nos deu a possibilidade de explorar a experiência de usuários, revelando problemas que aconteceram após o uso por um longo período. Por outro lado, ele não prevê especificamente a classificação dos problemas em relação às características como mudança de contexto, mobilidade, calma, entre outras.

Como resultado da avaliação, podemos dizer que o sistema *Hidrate Spark* possui muitos problemas de usabilidade. Esse fato é pior nos dispositivos Android. Muitos problemas de conectividade e de entendimento de como e quando ocorre a troca de informações entre a garrafa e aplicativo do *smartphone*, foram revelados, dando indício de que existe ainda uma grande dificuldade de as pessoas interagirem com sistemas ubíquos e de entenderem seu funcionamento. Por outro lado, na entrevista do Teste de Usabilidade, percebemos o interesse por usar algo inovador.

REFERÊNCIAS

1. Simone D. J. Barbosa and Bruno S. da Silva. 2010. *Interação humano-computador*. Elsevier.
2. Bruno Bonifácio, Priscila Fernandes, Fábio Santos, Horácio A. B. F. de Oliveira and Tayana Conte. 2012. Usabilidade de Aplicações Web Móvel: Avaliando uma Nova Abordagem de Inspeção através de Estudos Experimentais. In *Proceedings of the XV Iberoamerican Conference on Software Engineering (CIbSE 2012)*, 236-249.
3. Rainara M. Carvalho, Rossana M. C. Andrade and Káthia M. de Oliveira. 2018. AQUARIUM - A suite of software measures for HCI quality evaluation of ubiquitous mobile applications. *Journal of Systems and Software*, 136, 101-136.
4. Andrea Fontana and James H. Frey. 1994. Interviewing: The art of Science. In *The Handbook of Qualitative Research*, Norman K. Denzin and Yvonna S. Lincoln (eds.). Sage, Londres, UK, 361-376.
5. Marília S. Mendes. 2015. *MALTU - Um modelo para avaliação da interação em sistemas sociais a partir da linguagem textual do usuário*. PhD Thesis. Universidade Federal do Ceará, Fortaleza, BR.
6. Jakob Nielsen. 1993. *Usability Engineering*. Morgan Kaufmann.
7. Jakob Nielsen. 1994. Enhancing the Explanatory Power of Usability Heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)*, 152-158.
8. Jakob Nielsen. 1995. 10 Usability Heuristics for User Interface Design. Retrieved March 22, 2018 from <https://www.nngroup.com/articles/ten-usability-heuristics/>
9. Ronald Poppe, Rutger Rienks and Betsy van Dijk. 2007. Evaluating the Future of HCI: Challenges for the Evaluation of Emerging Applications. In *Artificial Intelligence for Human Computing*, Thomas S. Huang, Anton Nijholt, Maja Pantic and Alex Pentland (eds.). Springer, Heidelberg, Berlim, 234-250.
10. Larissa C. Rocha, Rossana M. C. Andrade, Andreia L. Sampaio and Valéria Lelli. 2017. Heuristics to Evaluate the Usability of Ubiquitous Systems. In *Proceedings of the 5th International Conference on Distributed, Ambient and Pervasive Interactions (DAPI 2017)*, 120-141.
11. Yvonne Rogers, Helen Sharp and Jenny Preece. 2011. *Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons.