

Análise de Métodos de Avaliação da Experiência do Usuário

Camilla V. L. T. Brennan Heiko Hornung
Cecília Baranauskas

Technical Report - IC-18-18 - Relatório Técnico
November - 2018 - Novembro

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE COMPUTAÇÃO

The contents of this report are the sole responsibility of the authors.
O conteúdo deste relatório é de única responsabilidade dos autores.

Análise de Métodos de Avaliação da Experiência do Usuário

Camilla V. L. T. Brennand¹, Heiko Hornung², Cecília Baranauskas³

Institute of Computing, University of Campinas (UNICAMP), Campinas-SP, Brazil
camillatenorio123@gmail.com; heiko@ic.unicamp.br; c.baranauskas@gmail.com

Abstract. Technology has become increasingly present in all spheres of society. However, the way we interact with technological devices has undergone significant changes. For example, in Internet of Things scenarios, functionality may be distributed on different devices with distinct capabilities; this distribution may cause interaction shifted in time and space. In this context, the human-computer interaction seeks to provide quality of interaction through measures coming from User Experience (UX) evaluations. This work presents an exploratory study using three UX evaluation methods: Attrakdiff, Emocards, and UX Curve, which include aspects such as temporality, emotion, hedonic and pragmatic characteristics. For the accomplishment of the study, a wearable device was used to emphasize the characteristics of IoT. The results showed gaps in the methods in the evaluation of UX in this scenario.

Keywords: User Experience, Methods evaluation User Experience.

1. Introdução

Este relatório descreve um estudo de UX na qualidade de análise exploratória de três métodos de avaliação de UX, considerando aspectos como temporalidade, emoções, características hedônicas e pragmáticas, com o propósito de analisar até que ponto estes métodos são adequados para avaliar experiências em contextos de Internet das Coisas ou *Internet of Things* (IoT) em inglês.

A primeira fase do estudo baseou-se na investigação de métodos com características relevantes para avaliação de UX no contexto citado. A segunda fase consistiu na seleção do objeto / sistema a ser utilizado no estudo. Foi usado como objeto de estudo a *smartband* SRW10 da Sony¹, um dispositivo vestível com funções de monitoramento de atividades físicas e de qualidade de sono, definição de metas e possibilidade de conexão com outros dispositivos para troca de informações. A escolha deste dispositivo se deu por suas características que compreendem atributos de objetos de IoT, como necessidade de conexão e distribuição de funcionalidades entre dispositivo e a aplicação.

Na terceira fase foram recrutados os participantes, convidados a participar do teste, fazendo uso do dispositivo por tempos distintos. Estes foram orientados a fazer uso do maior número de funções possível. O critério de escolha dos participantes, foi por proximidade e facilidade de acesso. No entanto, não foram determinados critérios, como escolaridade ou gênero. Foram recrutados cinco participantes que usaram a pulseira em média por oito dias cada um. Após o uso do dispositivo foram aplicados os métodos supracitados de maneira individual com cada participante. Na última fase foram analisados os resultados obtidos por meio da aplicação dos métodos, com enfoque nas dificuldades dos métodos em avaliar um cenário de IoT, bem como na fidelidade das respostas dadas pelos usuários.

Os resultados mostram algumas lacunas nos métodos usados no que diz respeito a avaliação da UX.

¹ <https://www.sonymobile.com/br/products/smart-products/>

1.1.O dispositivo

A “Sony SmartBand SWR10”¹ é um dispositivo vestível. O dispositivo deve ser usado no pulso e se conecta com um dispositivo Android que pode ser um telefone celular ou tablet. Através de um aplicativo (Lifelog)² os dados coletados pela pulseira são sincronizados, por exemplo, é possível verificar quanto tempo de sono profundo o usuário teve durante a noite, ou visualizar a quantidade de passos dados durante uma caminhada. Dentre outras, a *Smartband* possui funções de vibração que notifica sobre alarmes, além de mensagens e chamadas recebidas.

O dispositivo é composto por um bracelete de silicone e um hardware responsável pela coleta de dados e monitoramento. Os requisitos para funcionamento, são um sistema operacional Android, e conexão ativa com a Internet para transmissão dos dados.



Figura 1: Pulseira SmartBand (<https://support.sonymobile.com/br/swr10/>)

1.2.Métodos Usados

Os métodos foram escolhidos pelo seu reconhecimento na literatura e por suas características, que compreendem fatores que podem influenciar a UX mencionados anteriormente: emoções, temporalidade e qualidades hedônicas e pragmáticas.

O primeiro método usado foi o UX Curve, utilizado para avaliar como a experiência do usuário mudou ao longo do tempo. Este é um método de caneta e papel, que busca usar as memórias dos usuários para mensurar quanto a experiência se alterou conforme o tempo de uso.

O AttrakDiff é um questionário com 28 itens em escala semântica diferencial, para medir a percepção do usuário de um sistema. Traz vinte e oito itens de escala diferencial semântica. Sete itens capturam qualidade pragmática, quatorze itens capturam a qualidade hedônica e sete itens medem recurso.

² (<https://play.google.com/store/apps/details?id=com.sonymobile.lifelog>)

E por fim o EmoCards, um instrumento não-verbal que mede as respostas emocionais do usuário, este é composto por ilustrações faciais representando diferentes emoções. Cada face é a combinação de duas dimensões emocionais: prazer e estimulação.

1.3.Participantes

Cinco pessoas participaram do estudo de caso, quatro do sexo masculino e uma do sexo feminino. Todos os respondentes apresentados na Tabela 1, são pós-graduandos de mestrado ou doutorado. Três dos participantes não haviam feito uso de dispositivos vestíveis, como a pulseira. Dois participantes usaram um dispositivo parecido, um *smartwatch* por um período de aproximadamente cinco meses e o outro participante usou uma *smartband* por duas semanas. Todos os participantes demonstraram interesse em usar o dispositivo e suas funcionalidades.

Tabela 1: Participantes

Sujeitos	Idade	Escolaridade	Início	Tempo de Uso	
				Fim	Qtd Dias
S1	26	Doutorando (a)	30/ago	10/set	11
S2	25	Doutorando (a)	12/ago	22/set	10
S3	28	Doutorando (a)	21/set	02/out	11
S4	24	Mestrando (a)	10/out	17/out	7
S5	35	Doutorando (a)	18/out	26/out	8

1.4. Procedimento

Cada teste foi feito individualmente, sendo que, cada participante usou a pulseira por tempos distintos. Inicialmente os participantes foram orientados pelo avaliador a usar o dispositivo de maneira espontânea. No entanto, foi sugerido que os mesmos utilizassem o máximo de recursos possíveis. Cada participante fez uso da pulseira entre sete e onze dias.

1.5. A Avaliação

A avaliação foi realizada aplicando os métodos após a utilização do dispositivo, exceto o método UX Curve que foi aplicado durante o uso. O Emocards e o Attrakdiff foram aplicados em seguida.

2. Attrakdiff

Pessoas possuem e utilizam produtos interativos diariamente. Estes devem prover funcionalidades adequadas e prover usabilidade e outras qualidades de uso. Além de

usabilidade, as pessoas buscam estimulação³, a qual é definida pelo autor como: “sentimento de que você tem muita novidade e estímulo ao invés de se sentir entediado e subestimado pela vida.” (Hassenzahl 2004) e também a identidade.

- Estimulação: as pessoas se esforçam para o desenvolvimento pessoal, isto é, a melhoria de conhecimento e habilidades. Os produtos podem promover esse desenvolvimento, na medida que se tornam estimulantes.
- Identidade: as pessoas também se expressam através dos objetos (Prentice 1987). Eles querem ser percebidos de uma maneira específica. Um produto pode promover isso comunicando uma identidade desejada.

O AttrakDiff (Hassenzahl 2004) é um questionário que avalia aspectos hedônicos, estimulação, identidade e qualidades pragmáticas de produtos de software. O questionário foi originalmente feito em alemão, mas foi transcrito para o inglês.

- **Qualidade Pragmática**

Atributos pragmáticos são associados à facilidade com que um usuário consegue manipular o ambiente, o produto ou software. Se pensarmos pragmaticamente pensamos na funcionalidade e utilidade de um produto.

- **Estimulação Hedônica**

Os atributos ligados à estimulação hedônica estão relacionados ao fato das pessoas desejarem desenvolver suas habilidades e conhecimentos sobre o produto. Estes enfatizam o bem-estar psicológico e suportam os objetivos relacionados ao ser, como sentir-se competente ou conectado aos outros (Hassenzahl 2004).

- **Identificação Hedônica**

Os atributos relacionados à identidade hedônica são os que nos fazem identificar com o produto em um contexto social. Um exemplo disso seria um *site* onde você pode comunicar quem você é para o mundo exterior.

- **Atratividade**

Quando falamos sobre algo como atraente para nós, geralmente resumimos toda a experiência do produto. Avaliamos o produto como um todo e usamos palavras como bom, ruim, bonito e feio para descrever o produto. No AttrakDiff, a atração é usada para medir o apelo global de um produto e como estes afetam julgamento completo Hassenzahl (2003).

³ “Stimulation Feeling that you get plenty of novelty and stimulation rather than feeling bored and understimulated by life.” Hassenzahl (2003)

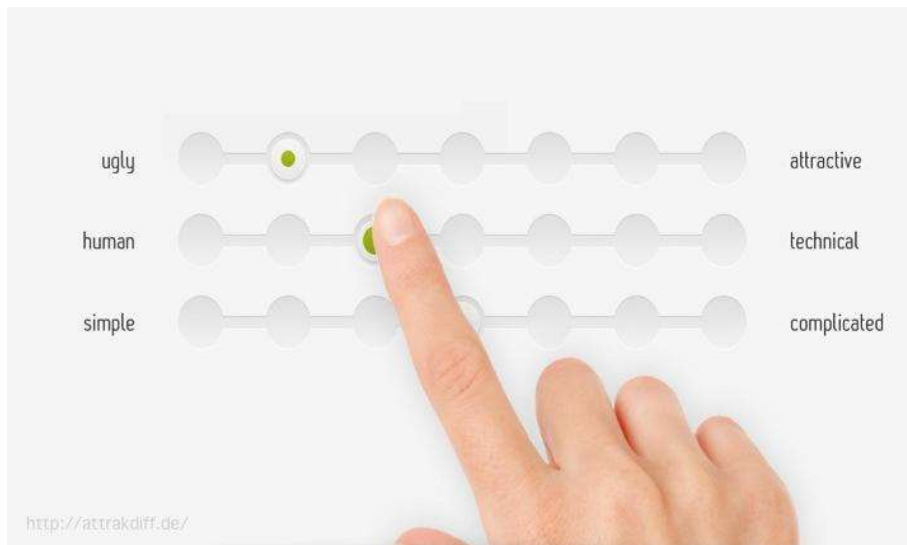


Figura 2: Exemplo Escala Attrakdiff (<http://attrakdiff.de/>)

2.1.O Método

O AttrakDiff consiste de um instrumento de medida no formato de diferenciais semânticos. É composto de 28 itens de sete passos cujos pólos são adjetivos opostos (por exemplo, "confuso - claro", "incomum - comum", "bom - ruim"). Cada conjunto de itens adjetivos é ordenado em uma escala de intensidade.

Cada um dos valores médios de um grupo de itens cria um valor para qualidade pragmática (PQ), qualidade hedônica (HQ) e atratividade (ATT).

A escala Qualidade Pragmática (PQ) tem sete itens cada com “bases” bipolares que medem as qualidades pragmáticas do produto. Isso inclui “bases”, como Técnico- Humano, Complicado-Simples, Confuso-Claro. HQI tem âncoras como Isolante-integrador, Barato-Valioso. HQS tem âncoras como Típico-Original, Cauteloso- Corajoso. A escala de Atratividade (ATT) com âncoras como feio-bonito e ruim-bom. As âncoras são apresentadas em uma escala, variando de 3 a 3, onde zero representa o valor neutro entre as duas âncoras da escala.

Attrak-Diff é um questionário para medir percepção do usuário sobre um sistema. Traz vinte e oito itens de escala de diferencial semântico. Sete itens capturam qualidade pragmática, catorze itens capturam qualidade hedônica e sete itens medem atratividade:

Qualidade pragmática

- Humanizado–Técnico
- Simples–Complicado
- Prático–Não prático
- Complicado–Claro
- Previsível–Imprevisível

- Confuso–Bem estruturado
- Desorganizado–Gerenciável

Qualidade hedônica

- Isolador–Conectivo
- Inventivo–Convencional
- Profissional–Não profissional
- Elegante–Desalinhado
- Inferior–Superior
- Alienador–Integrador
- Me aproxima das pessoas–Me afasta das pessoas
- Não apresentável–Apresentável
- Desimaginativo–Criativo
- Ousado–Cauteloso
- Inovador–Conservador
- Entediante–Cativante
- Pouco exigente–Desafiador
- Único–Comum

Atratividade

- Agradável–Desagradável
- Feio–Atraente
- Amigável–Não amigável
- Pouco atraente–Atraente
- Bom–Ruim
- Repulsivo–Sedutor
- Motivar–Desencorajador

2.2. Condução

- Para utilizar o questionário AttrakDiff, é necessário um cadastro prévio na plataforma. Uma vez realizado o cadastro, e *login* feito no site, é preciso criar um novo projeto com dados a respeito do produto, descrição de participantes e duração do projeto. Em seguida o questionário foi enviado via e-mail para os participantes juntamente com sua descrição.

- Os participantes responderam o questionário logo após o uso do dispositivo em um modelo baseado no AttrakDiff denominado Avaliação Individual, onde o produto é avaliado após o uso, e sem comparação com outro produto como é o caso da Comparação A – B.
- Foi indicado aos usuários que ao responder o questionário tentassem pensar em como o dispositivo e o aplicativo funcionam em conjunto, já que ambos fazem parte de um único sistema.
- Após obter as respostas do questionário os dados foram analisados de acordo com o procedimento do Attrakdiff.

2.3. Análise

O Attrakdiff trabalha com pares de palavras em um diferencial semântico de sete níveis. Tendo o ponto central como pontuação 0, os valores posicionados à direita são pontuados como 1, 2 e 3, e os valores pontuados à esquerda como -1, -2 e -3.

Os valores extremos segundo a metodologia devem ser analisados com cautela, pois estes mostram quais características são críticas ou bem consolidadas. Os resultados esperados trazem a maioria dos pontos posicionados à direita da linha divisória, ou seja, apresentando valores acima de 0 e, ou se possível, próximos de 3.

2.4. Resultados Attrakdiff

A seguir serão apresentados resultados obtidos da aplicação do questionário Attrakdiff. Tais resultados podem ser visualizados através da aba “Visão Geral dos Resultados”, estes podem ser exportados para posteriores análises.

Diagrama 1 - Descrição dos pares de palavras

Este diagrama apresenta os valores médios dos pares de palavras. Os valores extremos são de particular interesse.

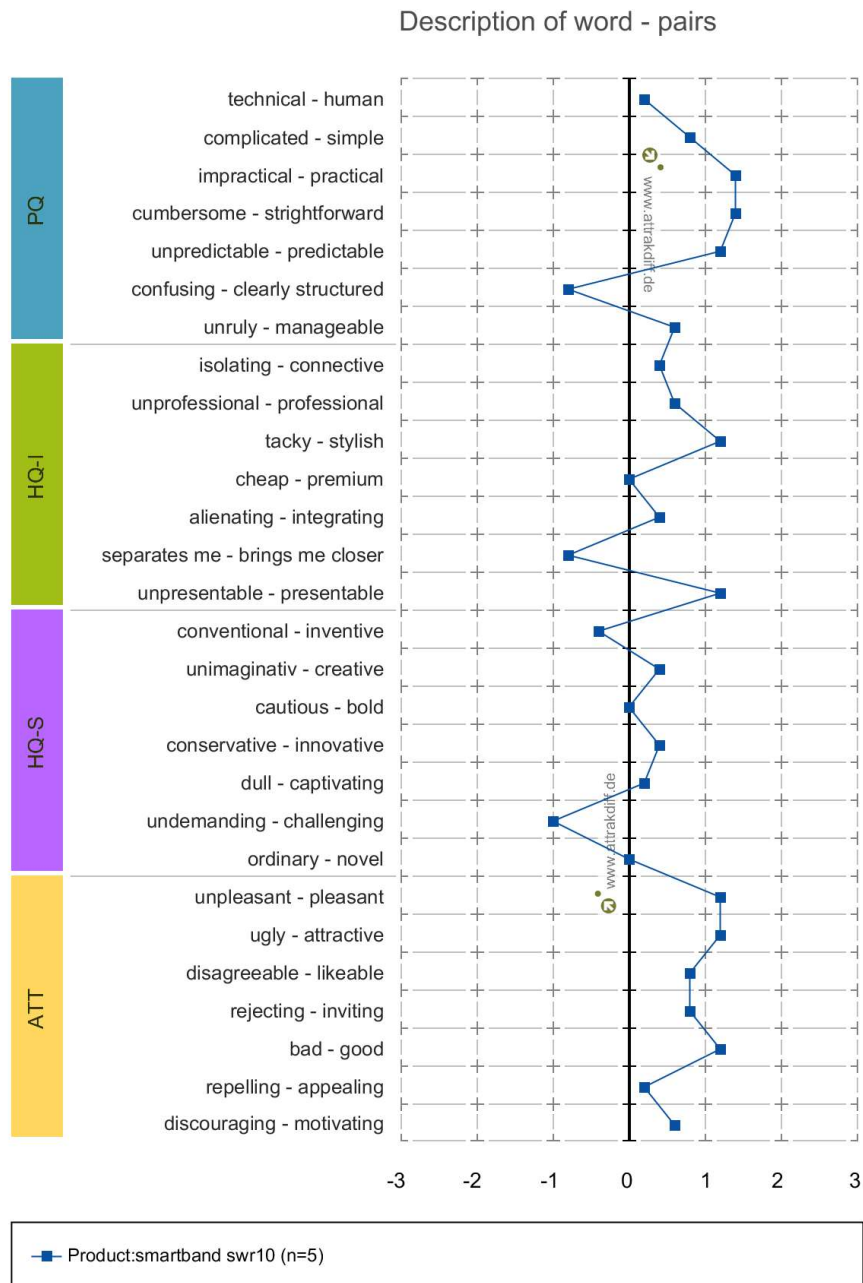


Figura 3: Resultado Descrição dos Pares de Palavras

Em relação à Experiência do Usuário, o diagrama apresentado na Figura 3 demonstra que a linha formada pela ligação dos pontos azuis está posicionada mais à direita, o que indica uma boa experiência em termos gerais. O gráfico mostra que a Qualidade Pragmática (PQ) obteve o maior número de resultados positivos, todavia o resultado para os atributos claro – confuso, apontou que o dispositivo/sistema é confuso.

Para Qualidade Hedônica - identidade (HQ-I), os itens Conectivo-Isolante, apesar de ser considerado conectivo, para os itens “me aproxima ou me afasta de”: o resultado foi negativo, mostrando que as opiniões são contrastantes, (ou seja, o dispositivo é conectivo, mas me afasta de algo).

A Qualidade Hedônica - Estímulo (HQ-S) apresentou itens posicionados no centro e a esquerda da linha central, tendo como destaque o adjetivo “despreocupado”. A percepção dos usuários sobre a atratividade (ATT), apresentou todos os itens posicionados à direita da linha evidenciando o adjetivo “agradável”. De maneira geral o dispositivo/aplicativo foi considerado atrativo.

No que se refere a forma de avaliação do método, com valores a direita pontuados em 1,2, e 3 a direita e a esquerda -1, -2, -3. Não houve nenhuma ocorrência de médias zero decorrentes de respostas positivas e negativas que carecesse de mais investigação.

O modelo de trabalho ilustra como as qualidades pragmáticas e hedônicas influenciam a percepção subjetiva de atratividade.

O modelo separa os quatro aspectos essenciais:

- A qualidade do produto projetada pelo designer.
- A percepção subjetiva de qualidade e avaliação subjetiva de qualidade.
- As qualidades pragmáticas e hedônicas independentes.
- Conspirações comportamentais e emocionais. Conforme mostrado na Figura 4.

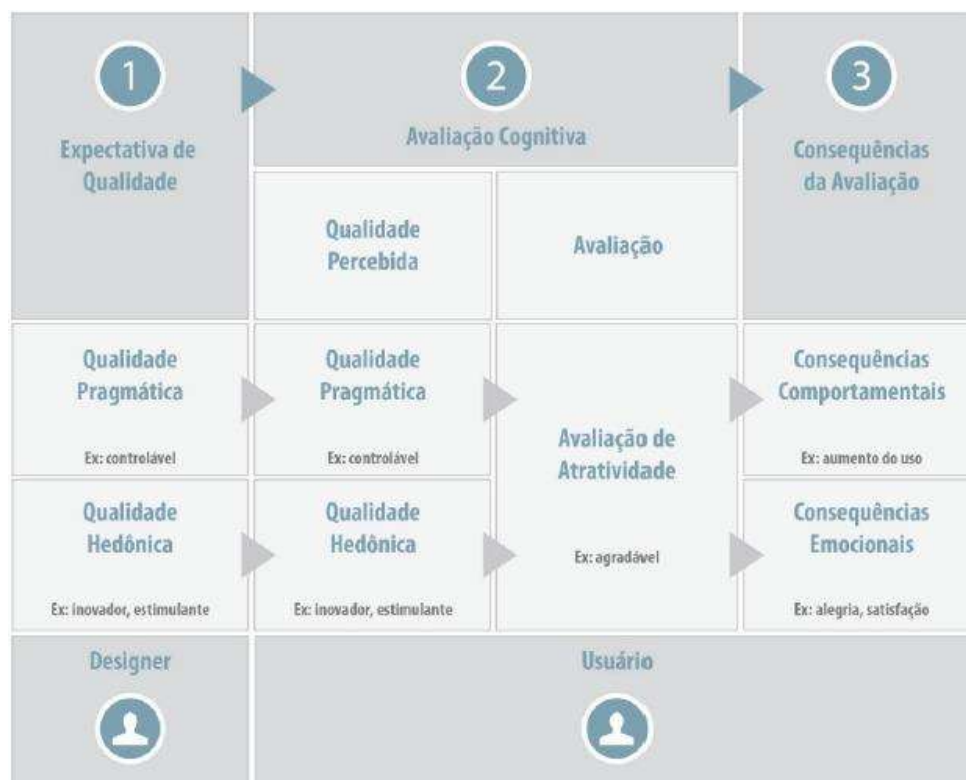


Figura 4 Modelo de Trabalho (<http://attrakidiff.de/>)

2.4.1. Portfólio de Resultados

O portfólio de resultados (Diagrama da Figura 5) mostra o posicionamento dos valores médios das dimensões: Qualidade Pragmática, e Qualidade Hedônica. A representação dos valores da Qualidade Hedônica é feita verticalmente, e os da Qualidade Pragmática horizontalmente.

Além disso, no diagrama são mostrados dois retângulos azuis, o menor e mais escuro representa o valor médio das dimensões do estudo com relação a experiência do usuário, o maior e mais claro representa o intervalo de confiança. O retângulo da confiança mostra o quanto as opiniões dos usuários convergem ou divergem. Quanto maior este retângulo, mais divergentes são as opiniões.

A experiência posicionou o retângulo que representa os valores médios no quadrante “neutro”, mais deslocado para a direita indicando percepção de qualidades pragmáticas (0,69) e hedônicas (0,19). O retângulo da confiança, apesar de localizado somente no quadrante neutro, mostra opiniões ligeiramente divergentes entre os participantes, tanto na dimensão Qualidade Pragmática (PQ), com valor de 0,30, quanto na Qualidade Hedônica (HQ), com valor de 0,54.

Portfolio-presentation

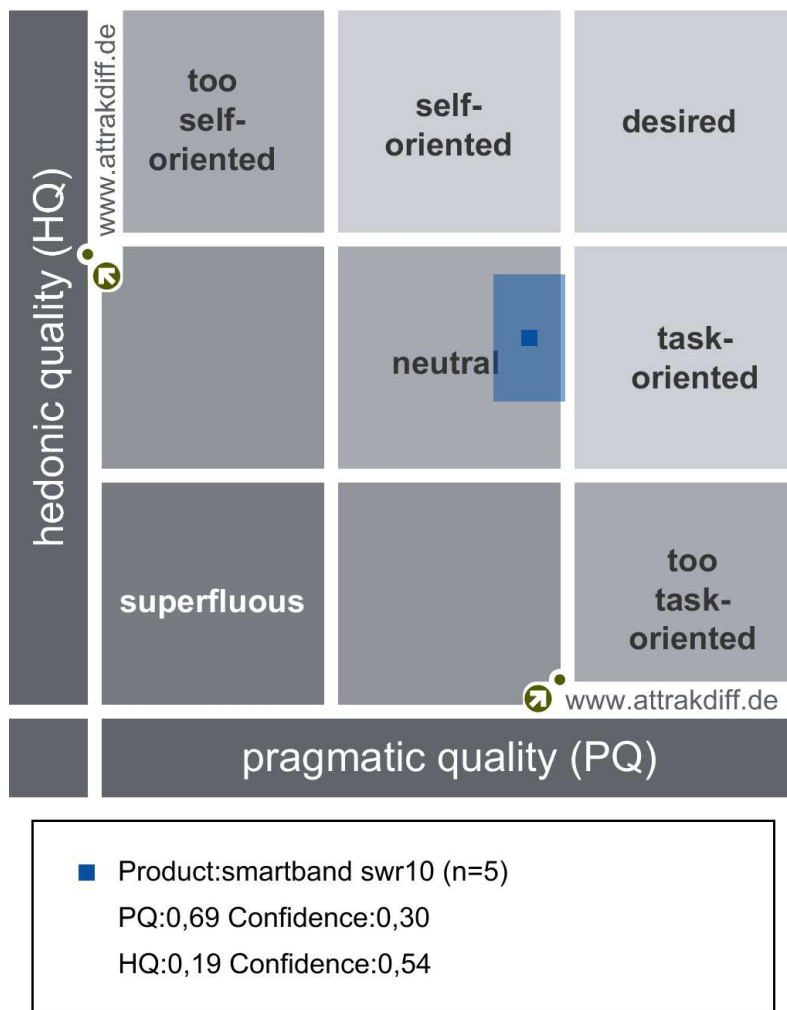


Figura 5: Portfólio Resultados

2.4.2. Diagrama de valores médios

O diagrama de valores médios (Diagrama da Figura 6), apresenta as dimensões do AttrakDiff para o produto. Nesta apresentação, a qualidade hedônica distingue entre aspectos de estimulação e identidade. O diagrama também apresenta a classificação de atratividade.

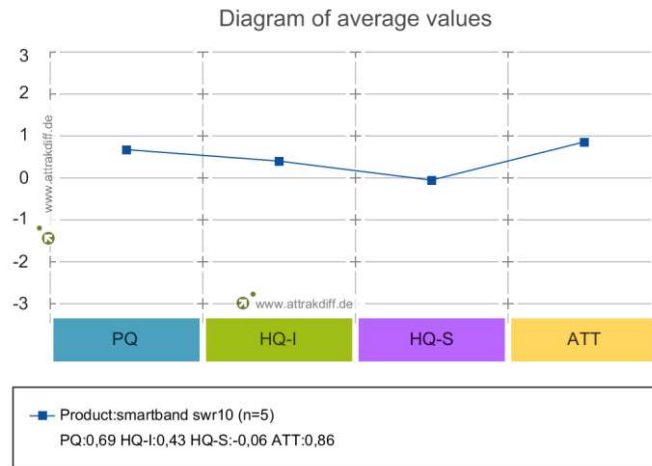


Figura 6: Diagrama Valores Médios

O diagrama de valores médios distribui os valores médios atribuídos pelos participantes nas dimensões Qualidade Pragmática (PQ), Qualidade Hedônica - Identidade (HQ-I), Qualidade Hedônica - Estimulo (HQ-S) e Atratividade (ATT).

A Qualidade Pragmática (PQ) indica o grau de sucesso no atingimento dos objetivos e obteve pontuação de 0,69 na experiência do usuário. O valor, acima de 0, indica que, de maneira geral, o resultado é positivo.

A Qualidade Hedônica - Identidade (HQ-I) indica o nível de identificação do leitor com o aplicativo e obteve pontuação 0,43, indicando que houve identificação.

A Qualidade Hedônica - Estimulo (HQ-S), que mensura se o aplicativo é original, interessante e estimulante obteve pontuação -0,06. O valor indica que o dispositivo/aplicativo não é estimulante.

A Atratividade (ATT), que indica quanto o aplicativo é atrativo para o usuário, obteve pontuação 0,86. Observando o (Diagrama 3) é possível notar que o resultado foi positivo, considerando que todos os itens ficaram posicionados à direita da linha central.

2.4.3. Discussão e Conclusão Attrakdiff

- Quando perguntados se foi fácil responder, os usuários mencionaram ter dificuldade com palavras cuja tradução não era conhecida;
- Uma das principais questões levantadas pelos participantes foi sobre qual parte do sistema avaliar, se o dispositivo, o aplicativo, ou o serviço gerado por ambos, o que pode ter levado a uma análise inconsistente, uma vez que estes mencionaram não saber a qual parte estaria destinada a questão;
- Apesar de bons resultados gerais, referentes ao dispositivo/aplicativo todos os participantes mencionaram que o mesmo não traz benefícios significantes que justifique

seu uso (ressaltando que todos usaram de forma voluntária); todos afirmaram que não comprariam o dispositivo;

- Ainda sobre os resultados positivos do questionário, os usuários disseram não perceber como o questionário avalia suas experiências, que as questões em grande parte tratam do produto analisado;
- Como mencionado anteriormente, a avaliação foi feita após o uso do produto, por parte de alguns participantes, algumas impressões segundo os mesmos podem ter sido perdidas no intervalo de uso.

Podemos observar através desta amostra que ficaram diversas lacunas. O fato do método segundo os participantes ser direcionado ao produto e não evidenciar a percepção sobre a experiência propriamente dita, mostra que ainda faltam aspectos a serem esclarecidos.

Em cenários de IoT comumente temos dois ou mais dispositivos que agem em conjunto, para prover um determinado serviço. Esta ligação entre dispositivos e serviço deve ser transparente ao usuário, a ponto de que o cenário como um todo represente benefícios desejados. Assim, uma questão relevante, no que diz respeito a avaliação de UX, trata-se da dificuldade dos participantes em associar estas partes ao responder o questionário.

Finalmente, embora os resultados para o dispositivo/aplicativo tenham sido positivos na média, os participantes afirmaram não ter interesse em uso posterior do dispositivo e afirmaram que o mesmo não traz benefícios.

3. Emocards

Todas as interações humanas envolvem emoções. As emoções são essencialmente pessoais: embora as pessoas possam diferir em suas respostas emocionais aos produtos, regras gerais podem ser identificadas no processo de emoção (Desmet 2001). Embora essas emoções sejam essenciais, é complexo saber sobre como as pessoas respondem emocionalmente aos produtos ou serviços e quais características desencadeiam reações emocionais positivas ou agradáveis. Além da dificuldade na compreensão das emoções e quais resultados elas provocam, é também difícil para os usuários discutirem suas respostas emocionais.

O EmoCards, é um instrumento não verbal que mede respostas emocionais do usuário, informadas por eles. Composto por 16 ilustrações faciais, metade remetendo o gênero masculino e metade ao feminino, representando diferentes emoções. Cada face é a combinação de duas dimensões emocionais: prazer e excitação. A premissa do Emocards é que as emoções desencadeadas sejam expressadas naturalmente pelos respondentes.

3.1.O Método

O método traz 16 figuras chamadas Emocards; os 16 emocards retratam rostos de desenhos com oito expressões emocionais distintas, oito rostos masculinos e oito femininos. Neste modelo, as emoções são descritas em termos de dimensões (Agradável-desagradável e excitante-calmo). As emoções variam entre muito agradável (por exemplo Emocionado) até

muito desagradável (por exemplo Horrorizado).. Destas oito categorias emocionais, as excitadas estão localizadas no lado superior do círculo, e as calmas no lado inferior. O lado esquerdo do círculo localiza emoções desagradáveis e o lado direito situa as agradáveis. Em síntese, os Emocards devem ajudar os sujeitos a expressar suas respostas emocionais, o usuário pode selecionar um cartão que melhor expresse sua resposta emocional a um produto ou serviço. A Figura 6 apresenta os Emocards, o grau de emoção proposto por Desmet (2001) e a descrição da emoção segundo Russell.

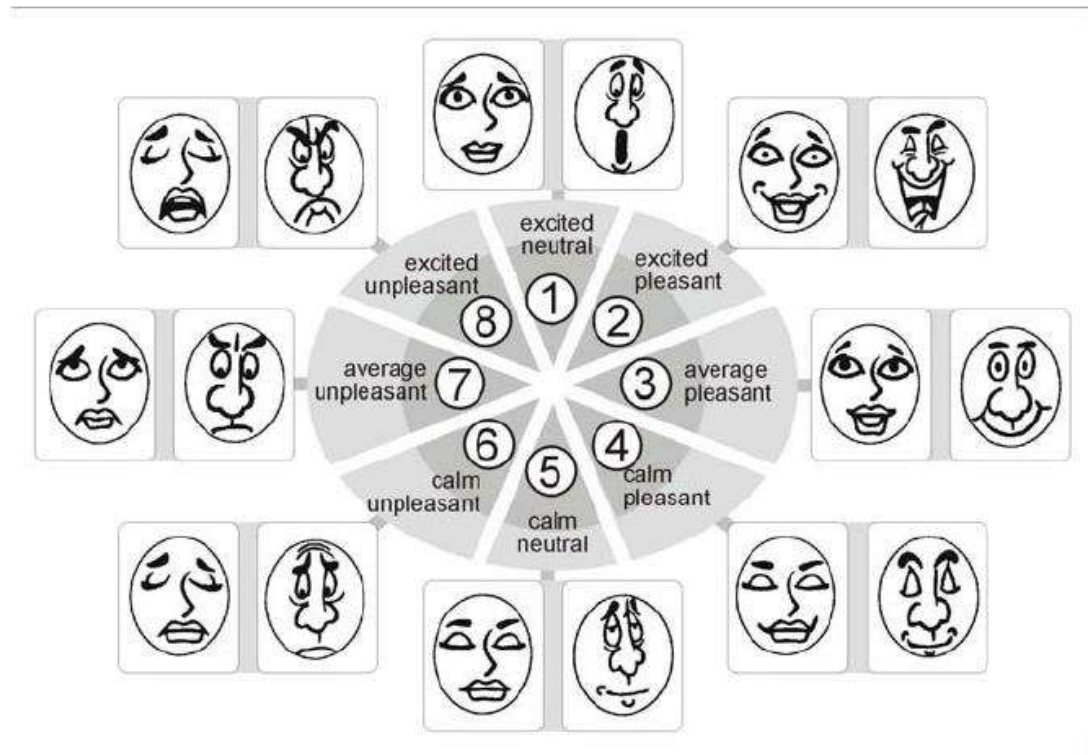


Figura 7: Modelo Emocards -Categorias Emocionais e Emocards Fonte Desmet (2001)

3.2. Metodologia

Após o uso do dispositivo os usuários foram convidados a avaliar o mesmo por meio do Emocard. As respostas emocionais ao produto/serviço e uso foram avaliadas de acordo com o procedimento proposto pelo método.

Neste estudo de caso, o pesquisador discutiu as preocupações subjacentes às respostas emocionais dos usuários em uma entrevista usando o método Emocards. As entrevistas foram realizadas individualmente. Foi escolhido um ambiente informal, para que os participantes se sentissem à vontade e pudessem se expressar livremente.

As entrevistas foram baseadas na técnica 'laddering' (Reynolds e Gutman 1988), assim como no método original Emocards. A característica básica de uma entrevista em escada é que a discussão começa com características concretas do produto e evolui para um nível mais abstrato. Desta forma, o tema da conversa evolui naturalmente. Esta técnica de entrevista em profundidade é baseada na repetição da pergunta: "por que isso é importante para você?" (vide subseção 3.2.1 Entrevista).

Para aplicar o método, as figuras foram apresentadas em ordem aleatória aos usuários, em seguida eles foram instruídos a selecionar o par de figuras que melhor representasse seu estado emocional em relação ao dispositivo e ao aplicativo, que haviam sido usados previamente. Em seguida, foi solicitado aos usuários que escolhessem qual o tipo de emoção eles sentiram, a fim de identificar quais tipos de emoções foram provocadas pelo uso produto. A avaliação foi dividida em duas partes, um questionário para o aplicativo e outro para o dispositivo, com intuito abranger a maior parte da experiência, cobrindo as partes do sistema de IoT.

O questionário baseou-se nas três etapas a seguir:

Passo um: O usuário deve mostrar no celular o aplicativo (através do qual os recursos da pulseira são acessados) e explicar quais deles ele aprecia.

Passo dois: O usuário é perguntado por que ele aprecia esses recursos; que efeitos/impressões agradáveis ele teve ao usar. Desta forma, a conversa é levada para um nível mais abstrato.

Passo três: Neste último passo da entrevista, o usuário é solicitado a explicar por que essas consequências são importantes para ele.

- Foram propostas duas entrevistas distintas, uma referente ao dispositivo e outra referente ao aplicativo, embora ambos façam parte de um único sistema.

3.2.1. Entrevista

- Pergunta 1: Quais recursos da pulseira você utilizou? (Pedir para mostrar)
 - 1.1 De quais recursos você mais gostou? (Para cada recurso pedir para o usuário apontar um emocard)
 - 1.1.1 Para cada recurso
 - 1.2 Por que você gostou deles?
- Pergunta 2: Quais recursos do aplicativo você usou?
 - 2.1 De quais recursos você mais gostou?
 - 2.2 Por que você gostou deles?

3.2.2. Recursos do Aplicativo (Lifelog)

*Com o objetivo de facilitar a aplicação da entrevista, os recursos foram nomeados de R1 à R6 para representar os recursos do aplicativo, na ordem em que estão aparecem na tela do aplicativo (Não necessariamente a ordem a serem usados).

- **R1** - Registros do que acontece em um determinado momento.

(Uma notificação aparece na tela do smartphone para que o usuário possa escrever uma nota sobre aquilo que deseja lembrar mais tarde).

- **R2** - Anotação das calorias queimadas, em exercício e passivas, os passos dados, as horas de caminhada, seguida pelas horas de corrida. (O que é calculado com base no GPS e nos dados de peso e altura que o usuário fornece assim que instala o aplicativo).
- **R3** - Contagem de horas de sono, com dados de quanto tempo de sono profundo, leve e acordado.
- **R4** - É possível contar o tempo gasto em redes sociais e comunicações em geral – Facebook, Instagram, Gmail e WhatsApp, por exemplo –, número de fotos tiradas e vistas, minutos de música escutado e tempo gasto em vídeos.
- **R5** - Na terceira e última linha aparecem as horas jogando (desde que conectado a PlayStation Network), de leitura e de pesquisa na internet.
- **R6** - Para cada atividade que o Lifelog registra é possível definir metas e que o aplicativo contabiliza os dados por dia, semana, mês e ano. Também é possível ver as distâncias percorridas em um mapa.

Sequência:

- Quais recursos do aplicativo você usou?
- De quais recursos você mais gostou
- Por que você gostou deles?

1. Quais recursos do aplicativo você utilizou? (Pedir para mostrar)

R1 – () **Emocard:**

R2 – () **Emocard:**

R3 – () **Emocard:**

R4 – () **Emocard:**

R5 – () **Emocard:**

R6 – () **Emocard:**

1.1 De quais recursos você mais gostou? (Para cada recurso pedir para o usuário apontar um emocard)

2.1.1 Para cada recurso

R1 – () **Emocard:**

R2 – () **Emocard:**

R3 – () Emocard:

R4 – () Emocard:

R5 – () Emocard:

R6 – () Emocard:

1.2 Por que você gostou deles? (Como se sentiu?)

R1 – () Emocard:

R2 – () Emocard:

R3 – () Emocard:

R4 – () Emocard:

R5 – () Emocard:

R6 – () Emocard:

2 De um modo geral como você classifica o aplicativo?

Emocard:

3.2.3. Recursos da Pulseira

*Com o objetivo de facilitar a aplicação da entrevista, os recursos foram nomeados como R1P, R3P e R4P os quais correspondem aos recursos do aplicativo e que podem ser alterados pelo botão do dispositivo.

- **R1P** - Ativação da função em que marca em determinado momento do dia a hora e temperatura local.
- **R3P** - Alterar modos noturno e diurno através do botão lateral.
- **R4P** - Notificações através de vibração da pulseira.

- **Questão 2:** Quais recursos da pulseira você utilizou? (Pedir para mostrar)

R1P – () Emocard:

R3P – () Emocard:

R4P – () Emocard:

2.1 De quais recursos você mais gostou? (Para cada recurso pedir para o usuário apontar um emocard)

2.1.1 Para cada recurso

R1P – () Emocard:

R3P – () Emocard:

R4P – () Emocard:

2.2 Por que você gostou deles? (Como se sentiu?)

R1P – () Emocard:

R3P – () Emocard:

R4P – () Emocard:

3. De um modo geral como você classifica o dispositivo?

Emocard:

Análise

Para analisar as respostas dadas pelos participantes, os Emocards foram graduados e divididos em valências: positiva, neutra e negativa. Esse critério possibilita a análise das respostas dadas pelos sujeitos de pesquisa para cada questão. A avaliação das valências utilizou como critério de avaliação, o proposto por Desmet (2001): os emocards graduados em zero, um e dois são considerados com valências negativas, os emocards três e quatro são considerados com valências neutras e os emocards cinco, seis e sete com valências positivas.

3.3.Resultados Emocards

Tabela Resultados por valências, na Tabela 1 são mostradas quantas vezes para cada recurso foram escolhidos os Emocards. Referente ao aplicativo.

Tabela 2: Repostas por Valências Aplicativo

Questões	Aplicativo																		2
	1						1.1						1.2						
Recursos	R1	R2	R3	R4	R5	R6	R1	R2	R3	R4	R5	R6	R1	R2	R3	R4	R5	R6	
Excited Neutral	1	1		2			1			1			1						1
Excited Pleasant	2	2		1	2	1	1	3	2	1	1			1	3	1	1	1	2
Average Pleasant	3	1	5	1	1	1	1		3				1	4	1				1
Calm Pleasant	4					1									1				
Calm Neutral	5	1		1		3	1	1				3	1				1	1	1
Calm Unpleasant	6																	1	
Average Unpleasant	7				1						1		1				1		
Excited Unplesant	8			2	1					1	1					2			
Qtd Respostas por Usuário																			

Questão 1: Quais recursos do aplicativo você usou?

- **Recurso 1:** Dois usuários responderam “Excitado Agradável”, um usuário optou por “Excitado Neutro”, um “Calmo Neutro”, um usuário “Médio Agradável”. Todas as respostas estão localizadas nas escalas positiva e neutra. Nestas escalas, os Emocards refletem emoções como, “Feliz”, “Animado”, “Contente”, “Serenos” e “Relaxado”.
- **Recurso 2:** Todos os usuários preferiram “Médio Agradável”, este resultado sugere uma relação entre atitude e resposta preferida.
- **Recurso 3:** Dois usuários indicaram preferir “Excitado Neutro”, um “Excitado agradável”, um “Médio agradável”, e um “Calmo Neutro”. Todas as respostas pertencem a escala positiva, denotando estados de felicidade e uma resposta neutra, deixando claro que o recurso 3 trouxe respostas emocionais positivas.
- **Recurso 4:** Dois usuários optaram por “Excitado Agradável”, um “Médio agradável”, dois “Excitado desagradável”. Metade dos resultados estão situados na escala positiva e o restante na escala negativa, o que denota uma divergência de opiniões entre os participantes.
- **Recurso 5:** Um usuário respondeu “Excitado Agradável”, um “Médio Agradável”, um “Médio Desagradável”, “Excitado Desagradável”. Novamente os usuários demonstram divergir em suas opiniões.
- **Recurso 6:** um “Excitado Agradável”, um “Calmo Agradável” e três responderam “Calmo Neutro”. As respostas para o recurso 6 apontam em sua maioria para escala neutra, com emoções como “Serenos” “Relaxado” e “Calmo”.

Questão 1.1 De quais recursos você mais gostou?

- **Recurso 1:** Um usuário optou por “Excitado Neutro”, um “Excitado Agradável”, um “Média Agradável” e um “Calmo Neutro”. Para o recurso 1 (colocar o recurso) todos os participantes demonstraram sentir emoções situadas nas escalas positiva e neutra.
- **Recurso 2:** Três usuários preferiram “Excitado Neutro” e um “Calmo Neutro”. Todas as respostas fazem parte da escala neutra, representando emoções como “Contentamento” e “Serenidade”.

- **Recurso 3:** Dois participantes optaram por “Excitado Agradável” e três “Médio Agradável”. Todas as respostas estão situadas na escala positiva, denotando sentimentos de “Alegria”, “Animação” e “Exaltação”.
- **Recurso 4:** Um usuário preferiu “Excitado Neutro” e um “Excitado Agradável” e um “Excitado Desagradável”, para o recurso 4 (colocar recurso 4), houveram divergências entre os participantes, além da pergunta ser seletiva, (preferido), dois dos cinco participantes não responderam.
- **Recurso 5:** Um usuário escolheu “Excitado Agradável”, um “Médio Desagradável” e um “Excitado Desagradável”, para o recurso 5 dois dos três respondentes demonstraram sentimentos da escala negativa, como “Tensão”, “Nervosismo” e “Chateação”.
- **Recurso 6:** Apenas três usuários responderam esta questão, optando por “Calmo Neutro”, representando emoções como “Serenidade” e “Calma”.

Questão 1.2: Por que você gostou deles?

- **Recurso 1:** Um participante respondeu “Excitado Neutro” um “Médio Agradável”, um “Calmo Neutro” e um “Médio Desagradável”. Nesta questão, dois respondentes na escala neutra e os restantes nas escalas negativas e positivas, mostrando opiniões distintas com relação ao recurso.
- **Recurso 2:** Neste recurso quatro usuários responderam “Médio Agradável”, um “Excitado Agradável”, todos na escala positiva, denotando emoções como “Felicidade” e “Animação”.
- **Recurso 3:** Três usuários escolheram “Excitado Agradável”, um “Médio Agradável” e um “Calmo Agradável”. Novamente todas as respostas pertencem as escalas positiva e neutra.
- **Recurso 4:** Um participante excitado agradável, dois “Excitado Desagradável”, três dos respondentes apontaram emoções desagradáveis.
- **Recurso 5:** Um usuário apontou “Excitado Agradável”, um “Calmo Neutro”, e um “Médio Desagradável”.
- **Recurso 6:** Um usuário apontou “Excitado Agradável”, um “Calmo Neutro”, e um “Calmo Desagradável”.

Questão 2: De um modo geral como você classifica o aplicativo?

Quando perguntados sobre a impressão geral do aplicativo, um usuário respondeu “Excitado Neutro”, dois disseram “Excitado Agradável”, um “Médio Agradável” e um respondeu “Calmo Neutro”. Todas as respostas estão nas escalas positiva e neutra, denotando que a impressão geral, trouxe respostas como “Felicidade”, “Calma”, “Contente”, “Animado” e “Relaxado”.

A Tabela 3 apresenta as valências separadas por escalas, positiva, neutra e negativa. Por exemplo, o Recurso 1 obteve quatro valências pertencentes à escala positiva e uma valência pertencente à escala neutra, que foram especificadas previamente.

Tabela 3: Resultados por Escala Aplicativo

	Aplicativo																		Total	
	Questões																			
	1						1.1						1.2							2
Recursos	R1	R2	R3	R4	R5	R6	R1	R2	R3	R4	R5	R6	R1	R2	R3	R4	R5	R6		
Positivo	4	5	4	3	2	1	4	3	5	2	1		2	5	4	1	1	1	4	52
Neutro	1	0	1	0	2	4	1	1	0	1		3	1	0	1		1	1	1	19
Negativo	0	0	0	2	0	0	0	0	0		2		1	0	0	2	1	1	0	9
	Escala																			

3.3.1. Referente ao dispositivo

A Tabela 3 mostra resultados por valências; na Tabela 4 são mostradas quantas vezes para cada recurso foram escolhidos os Emocards, referentes ao aplicativo.

Tabela 4: Repostas por Valências Dispositivo

Questões	Pulseira									3
	2			2.1			2.2			
Recursos	R1P	R3P	R4P	R1P	R3P	R4P	R1P	R3P	R4P	
Excited Neutral 1										1
Excited Pleasant 2										
Average Pleasant 3		1	2				4			3
Calm Pleasant 4		1	2					1	1	
Calm Neutral 5	1			1	1		2	1		1
Calm Unpleasant 6			1							1
Average Unpleasant 7		1								
Excited Unplesant 8	2	1						1		
	Qtd Respostas por Usuário									

Questão 2: Quais recursos da pulseira você utilizou?

- **Recurso 1P:** Um “Calmo Neutro”, e dois “Excitado Desagradável”.
- **Recurso 3P:** Um “Médio Agradável”, um Calmo Agradável, um Médio Desagradável e um Excitado Desagradável. Em geral este recurso foi avaliado negativamente.
- **Recurso 4P:** Dois “Calmo Agradável”, dois “Médio Agradável e um “Calmo Desagradável”.

Questão 2.1: De quais recursos você mais gostou?

- **Recurso 1P:** Neste recurso apenas um usuário apontou “Calmo Neutro”.
- **Recurso 3P:** Um usuário “Calmo Neutro” e um “Calmo Agradável”. Com emoções como “Felicidade”, “Calma, Serenidade”.
- **Recurso 4P:** Neste recurso quatro dos cinco usuários responderam e optaram por “Médio Agradável”.

Questão 2.2: Por que você gostou deles?

- **Recurso 1P:** Dois “Calmo Neutro”.
- **Recurso 3P:** Um “Calmo Neutro”, um “Calmo Agradável”, um “Excitado Desagradável”. Mostrando divergências entre as impressões dos participantes.
- **Recurso 4P:** Três respostas “Médio Agradável”, um “Calmo Agradável” e um “Calmo Desagradável”. Idem ao item anterior.

Questão 3: De um modo geral como você classifica o dispositivo?

Quando questionados sobre sua impressão sobre o dispositivo, no diz respeito a sua aparência, conforto, etc. um participante mencionou “Excitado Neutro” um “Calmo Neutro” e três apontaram “Calmo Desagradável”. Apesar das respostas a questões individuais avaliarem positivamente o dispositivo, no que diz respeito a sua estética, este foi avaliado negativamente.

A Tabela 5 apresenta as valências separadas por escalas, positiva, neutra e negativa. Por exemplo, o Recurso 1 obteve quatro valências pertencentes a escala positiva e uma valência pertencente a escala neutra.

Tabela 5: Resultados por Escala Dispositivo

	Pulseira										Total
	Questões										
	2			2.1			2.2			3	
Recursos	R1P	R3P	R4P	R1P	R3P	R4P	R1P	R3P	R4P		
Positivo		1	2			4			3	1	10
Neutro	1	1	2	1	1	1	2	2	1	1	13
Negativo	2	2	1					1	1	3	10
	Escala										

3.3.2. Avaliação Geral

Além da entrevista estruturada, no final foi perguntado aos participantes a opinião deles a respeito do método. Lhes foi perguntado: “Você acredita que o método usado foi capaz de avaliar sua experiência com o produto? ”. Além disso, foi permitido que durante o teste os usuários fizessem observações, a respeito do dispositivo, serviço, método, etc. Estas observações estão apresentadas por usuário. As observações não interferiram nos resultados do método, mas foram levadas em consideração para análise final, como uma tentativa de avaliar o método.

Usuário 1:

Observações gerais:

- Com relação ao Recurso R4 “É possível contar o tempo gasto em redes sociais e comunicações em geral”, considera invasivo.
- Gostou das “carinhas”, mas não acha que elas representam corretamente as emoções.
- Não existe notificação quando se ativa o modo noturno.

Questão sobre o método: acredita que não avalia a experiência, mas sim o produto, as emoções ou sentimentos despertados são difíceis de serem lembrados, já que fez uso durante alguns dias e a avaliação é feita somente no final.

Usuário 2:

Observações gerais:

- O dispositivo é invasivo;
- Funções de vibração apesar de serem úteis, incomodam;
- Não pagaria o preço pelo dispositivo;

Questão sobre o método: Muito subjetivo, não entende quais emoções estão relacionadas a quais Emocards, gerando confusão. Apesar de ser bom expressar-se por imagens elas não refletem as emoções sentidas. Não faz sentido para o serviço, e para questões técnicas também é difícil apontar uma emoção.

Usuário 3

Observações gerais:

- “As saídas da pulseira são ruins e confusas”;

Questão sobre o método: “Não ter que escrever é bom, mas os Emocards masculinos e femininos não condizem com faces femininas com rostos femininos e masculinos. Nas questões sobre o “Porquê” não é possível externar as emoções através dos Emocards. Além disso seriam necessárias mais representações”.

Usuário 4

Observações gerais:

- Frustração por que alguns recursos não funcionaram;
- Com relação a necessidade de conexão, não viu relevância, não causaria danos se não conseguisse acessar, por outro lado, isto geraria uma frustração desnecessária.
- Difícil configuração e pareamento;
- Não compraria o dispositivo pois não agrega nada, é apenas mais um aparelho para interagir.

Questão sobre o método: É mais fácil expressar-se através das imagens, é fácil de entender. No entanto é difícil de lembrar das sensações para cada recurso.

Usuário 5

Observações gerais:

- Não funciona corretamente;
- Desconfortável;
- As notificações incomodam;
- O dispositivo nada acrescenta.

Questão sobre o método: É mais fácil expressar com figuras, avalia o produto e não como a pessoa se sente. Algumas faces não representam o sentimento. Deveria ter faces intermediárias. Não se lembra de alguns recursos.

3.3.3. Discussão e Conclusão Emocards

Com base nas análises efetuadas, foi possível perceber que o ambiente teve maior quantidade de valências positivas. Para Questão 1, onde foi solicitado que o usuário apontasse quais recursos utilizou, por exemplo o Recurso 2 (Anotação das calorias queimadas e contagem de passos), obteve de todos os usuários uma avaliação positiva.

Na Questão 1.1 “De quais recursos você mais gostou? ” O Recurso 2 citado anteriormente e Recurso 3 (Contagem de horas de sono) obtiveram maior preferência, sendo assinalados por todos os participantes. Alguns recursos não obtiveram respostas, já que era uma questão de preferência. Alguns recursos não foram usados pelos participantes, no entanto, três dos cinco participantes decidiram por si mesmos, imaginar a experiência que seria sentida se tivessem usado, estas respostas foram levadas em consideração para a análise final.

Todas as questões foram baseadas na técnica de entrevistas *laddering*, descrita anteriormente. Com ela foi possível lidar com o nível maior de subjetividade.

A Questão 1.2, que indaga “Por que você gostou do recurso”, tenta alcançar os níveis mais abstratos, o objetivo é ir além das questões objetivas. Idem a Questão 1.1 alguns recursos não foram usados. Os resultados da questão foram na maioria positivos para todos os recursos, exceto para o Recurso 4 (É possível contar o tempo gasto em redes sociais e comunicações em geral) que foi avaliada negativamente. Na escala negativa a valência “Excitado Desagradável” com emoções como “Nervosismo”, “Chateação” ou “Tensão” para o Recurso 4. Os resultados para a pergunta sobre a impressão geral, quatro dos cinco usuários, avaliaram positivamente.

Com relação aos resultados do dispositivo, que se referiam a recursos físicos, (como por exemplo a funcionalidade na qual altera-se de modo noturno para o diurno e vice-versa através do botão lateral), estes demonstraram uma equivalência, as avaliações positivas e negativas alcançaram os mesmos resultados e as valências neutras um resultado um pouco melhor. Para a Questão 2 “Quais recursos da pulseira você utilizou? ”, para o Recurso 1P “Ativação da função que marca em determinado momento do dia a hora e temperatura local”, dois dos respondentes descreveram como “Excitado Desagradável” e um “Calmo Neutro”. Para o Recurso 3P (“Alterar modos noturno e diurno através do botão lateral”), ocorreram respostas nas três escalas. O Recurso R4P (Notificações através de vibração da pulseira) obteve maior aprovação.

O questionamento sobre impressão geral também trouxe resultados nas três escalas, todavia com um número maior de respostas negativas.

A tabulação permitiu inicialmente uma visão geral do ambiente, o qual foi caracterizado pelos sujeitos de pesquisa como positivo, sendo “Felicidade”, “Exaltação” e “Animação” pertencentes à escala positiva, e “Serenidade” “Calma” e “Contentamento” da escala neutra mais assinalados. Todavia, a avaliação do dispositivo físico trouxe mais valências negativas, embora não tenha sido avaliado negativamente na totalidade.

Além dos dados obtidos através do método, foi permitido que os participantes se expressassem fazendo observações. Estas observações trouxeram informação útil sobre a premissa inicial, que busca compreender a capacidade do método em avaliar a UX como um todo.

Apesar dos resultados positivos apontados pelo método no que diz respeito à experiência, é possível concluir que, para avaliar a experiência do usuário, é preciso cobrir todos os aspectos da interação, considerando o usuário, o contexto, o sistema como um todo.

Um outro ponto importante, são os resultados da questão sobre a percepção do usuário a respeito da avaliação da UX utilizando o Emocards. Foi unânime a posição dos participantes a respeito da facilidade de apontar figuras ao invés de escrever a emoção sentida. Do mesmo modo, foi consensual a opinião sobre os Emocards não representarem as emoções sentidas, e a dificuldade em diferenciar as figuras.

4. UX curve

Tradicionalmente, métodos que estudam a UX focam em experiências momentâneas, que ocorrem na interação com produtos ou serviços na primeira vez. No entanto, houve um crescente interesse no aspecto da temporalidade ao estudar a experiência do usuário (Karapanos 2009).

Segundo Kujala (2011), avaliar a experiência momentânea é na maioria dos casos menos confiável do que a avaliação a longo prazo. Tendo em vista a relevância de estudar a UX a longo prazo, o UX curve foi usado na avaliação da *smartband* Sony. Uma das razões para escolha do método é a facilidade de aplicação na avaliação e o baixo custo.

4.1. O Método

O UX Curve é um método que utiliza caneta e papel. O participante desenha uma ou mais curvas para descrever como a experiência sobre um produto ou serviço mudou ao longo do tempo. A área de desenho da curva é formada por uma linha de tempo e uma linha horizontal que divide experiências positivas e negativas (Kujala 2011).

As instruções fornecidas pelo modelo são: "Lembre-se do momento em que você começou a usar o produto. Em seguida, desenhe uma curva descrevendo como seu relacionamento com o produto mudou desde a primeira vez que você usou até hoje. O modelo compreende uma área de gráfico bidimensional vazia e linhas para descrever sucintamente os motivos das mudanças na curva. O eixo horizontal representa a dimensão tempo desde o início do uso até o momento atual e o eixo vertical representa a intensidade da experiência dos usuários. Na região do meio do gráfico existe uma linha horizontal de (valor zero) dividindo a área em uma parte superior positiva e uma parte inferior negativa. O eixo vertical é rotulado de acordo com o sinal (+) para parte superior e (-) para a parte inferior. A Figura 8 mostra o modelo usado para avaliação do UX Curve.

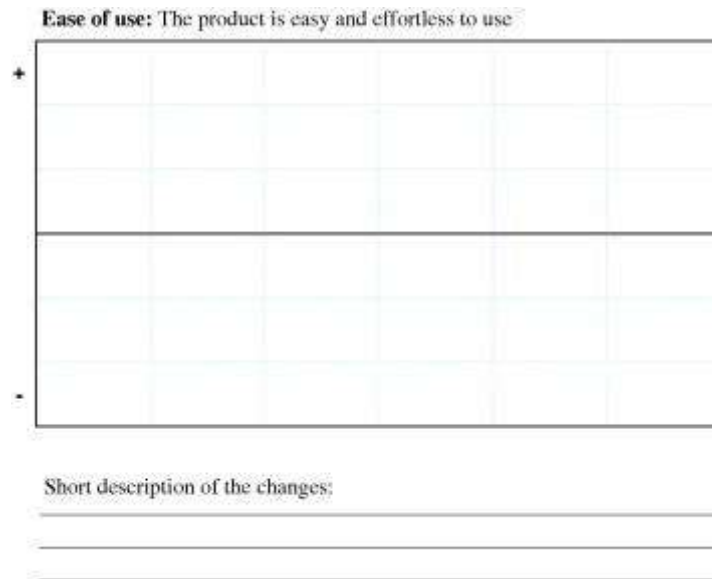


Figura 8: UX Curve Template - Fonte Kujala (2011)

4.2. Metodologia

As sessões com os participantes consistiram na descrição do método, em seguida iniciou-se a sessão de desenho de curva, incluindo descrições dos motivos das mudanças em suas experiências.

Os usuários participaram individualmente das sessões, que não tinham duração máxima. Primeiro, foi pedido ao participante que recordasse o momento em que eles começaram a usar o dispositivo. Foi entregue um formulário do UX Curve vazio, para a curva UX geral e os participantes foram aconselhados a desenhar uma curva descrevendo como a relação deles com o dispositivo/aplicativo mudou desde a primeira vez que o usaram até o presente. Os participantes poderiam escrever os motivos na área abaixo do gráfico se desejassem. No modelo geral de Curva UX, o usuário pode expressar-se livremente.

Os motivos que explicam as mudanças da tendência da curva, mencionadas durante a sessão e escritas no modelo da curva, foram analisados partir do documento de texto.

Finalmente, o pesquisador perguntou aos participantes, para obter feedback sobre o método: “Você acredita que este método conseguiu avaliar sua experiência ao longo do tempo? ”.

4.3. Análise de Dados

Os dados foram analisados de acordo com a definição do método. As curvas abaixo da linha foram consideradas negativas, as próximas a linha central consideradas neutras e as curvas desenhadas acima da linha foram consideradas positivas. Foram considerados os picos das curvas, os motivos das alterações e as descrições de quando estas foram feitas, abaixo do gráfico. As anotações foram avaliadas com base no modelo de experiência do usuário de Hassenzahl (2003). O modelo categoriza problemas de experiência do usuário em pragmática (utilidade e usabilidade) e qualidade hedônica ou não instrumental.

Além da classificação positiva, neutra e negativa, também foi feita a categorização com base no ponto de partida da curva, caso esta tenha sido maior ou menor em comparação com o ponto final. Por exemplo, a curva foi categorizada como sendo *melhorada* quando o ponto inicial era inferior ao ponto final, ainda que a curva se alterasse durante o “trajeto”. Se os pontos inicial e final estivessem no mesmo nível, a curva seria categorizada como *estável*. Como as curvas são desenhos à mão livre, eles foram categorizados como estáveis se o desvio foi muito pequeno (menos de um milímetro) entre os valores verticais dos pontos inicial e final da curva. Caso o ponto inicial da curva estivesse acima do ponto final, esta foi classificada como *deteriorada*. Os dados coletados foram analisados com análise estatística simples, já que a amostra foi pequena.

Tabela 6: Resultados Picos por Usuário

		Motivos	Classificação
Usuário 1	1º Pico	Gráfico legal	Positivo
		Não incomoda	Positivo
	2º Pico	Não mede batimentos cardíacos	Negativo
		Incomoda para vestir casacos	Negativo
	3º Pico	Gostou do monitoramento de sono	Positivo
4º Pico	Revela tempo gasto em redes sociais	Positivo	
Usuário 2	1º Pico	Empolgado com a pulseira	Positivo
	2º Pico	Notou falhas técnicas	Negativo
	3º Pico	Aprendeu novas funcionalidades	Positivo
	4º Pico	Perdeu o interesse (falta de recursos e incômodo)	Negativo
Usuário 3	1º Pico	Acreditou que seria legal	Positivo
	2º Pico	Nada funcionou	Negativo
	3º Pico	Várias funções não funcionavam	Negativo
	4º Pico	Notificações inconvenientes	Neutro
	5º Pico	É bacana mas não compraria	Neutro
Usuário 4	1º Pico	Difícil conexão com celular	Negativo
	2º Pico	Manual do usuário	Neutro
	3º Pico	Uso do aplicativo	Positivo
	4º Pico	Uso das funcionalidades da smartband	Positivo
Usuário 5	1º Pico	Bonita	Positivo
	2º Pico	Conta passos	Positivo
	3º Pico	Mede sono	Positivo
	4º Pico	Desconfortável para dormir	Negativo
	5º Pico	Conta passo errado	Negativo

Os resultados foram classificados com base nos picos (onde a curva alterou-se) e os usuários fizeram uma breve descrição. Foram coletados um total de vinte e dois (22) picos com uma média de quatro picos para cada usuário. Do total foram onze picos positivos com comentários do tipo, “O gráfico é legal (apresentado na primeira tela com informações de atividades)”, “Estou empolgado com o uso” e a “Contagem de passos é legal”. Três ocorrências foram consideradas neutras, por estarem localizadas a uma distância mínima da linha central. No entanto, apesar de classificadas como neutras em virtude da localização no gráfico, as justificativas poderiam ser catalogadas como negativas, por exemplo a frase “É bacana, mas não compraria”. Oito das ocorrências foram negativas, com comentários como: “Incomoda para dormir” ou “Difícil conexão com celular”.

4.3.1. Tendências das Curvas

A Figura 9 mostra um exemplo de curva UX geral desenhada pelo usuário. Estas curvas mostram uma tendência típica. Uma expectativa maior ao iniciar o uso e um decréscimo após frustrações relacionadas a funcionalidades, por exemplo.

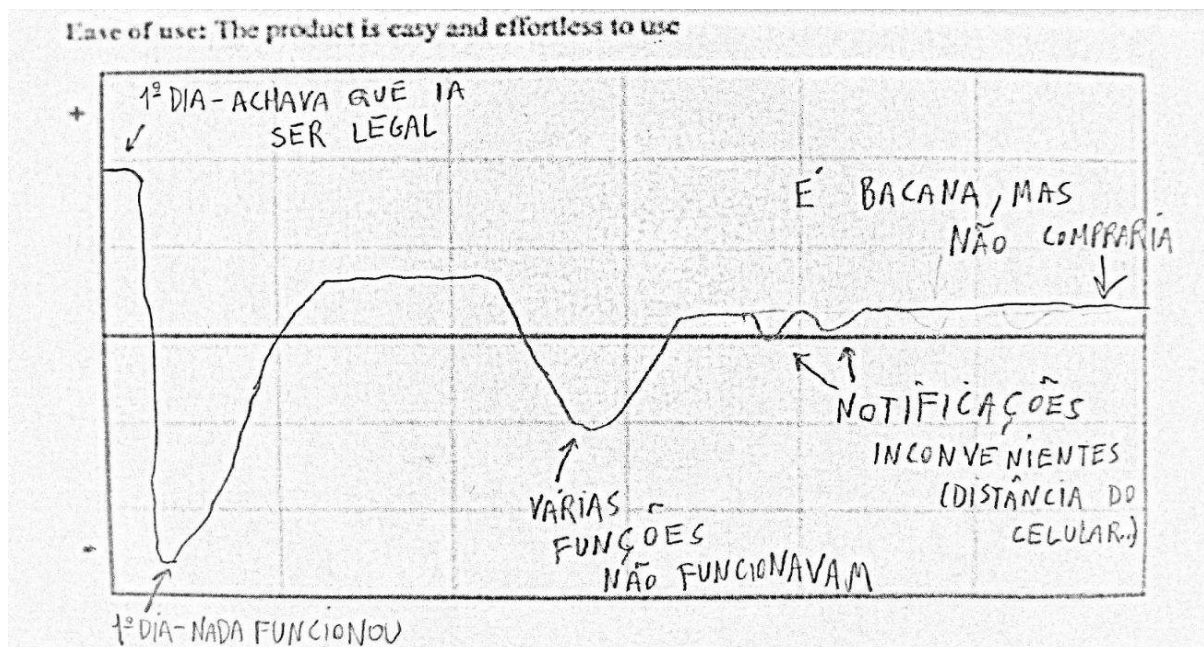


Figura 9: Exemplo de curva UX

Uma outra tendência mostrada na Figura 10, é o crescimento da curva quando o usuário começa a alcançar suas expectativas; por exemplo, quando aprende a usar uma nova funcionalidade ou um recurso desejado.

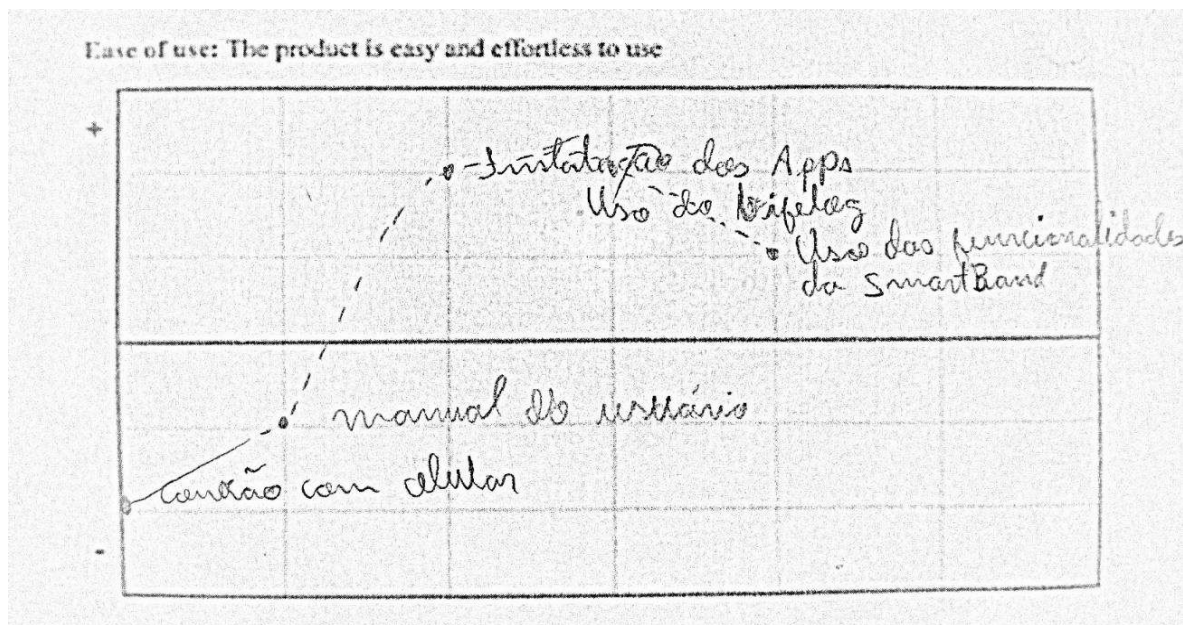


Figura 10: Exemplo do crescimento da curva UX

4.3.2. Análise Tendências

A Figura 11 mostra uma reprodução das curvas coletadas durante o teste. A maioria das curvas inicia com empolgação ou melhores expectativas e uma diminuição da qualidade da experiência por assim se dizer. Por exemplo, os usuários U2, U3 e U5 demonstram este declínio, esta deterioração está relacionada a dificuldades causadas por defeitos técnicos ou características frustrantes individuais. No entanto, os usuários U1 e U4 representam uma outra tendência comum, quando se analisa a UX ao longo do tempo, o usuário começa com pouca expectativa, mas esta muda conforme ele usa ou aprende novos recursos. Não houve respostas que demonstrassem a curva Neutra, o que pode dever-se ao tamanho da amostra. A qualidade dos motivos relacionados com as curvas de melhoria ou deterioração foi bastante semelhante. Falhas técnicas e erros foram mencionados mais frequentemente com relação às curvas deterioradas do que as curvas melhoradas. Os usuários com curvas deterioradas relataram mais razões em média, do que os usuários com curvas melhoradas.

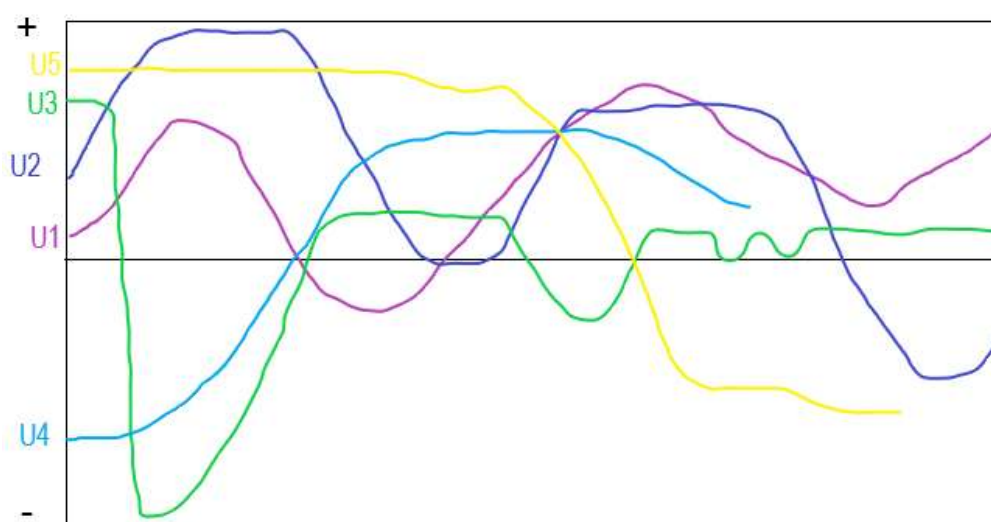


Figura 11: Curvas Coletadas

4.3.3. Razões para a Mudança de Experiência do Usuário

Além de desenhar as curvas, os usuários descreveram os fatores que melhoraram sua experiência ao longo do tempo ou causaram deterioração. Os motivos foram variados, problemas práticos, expectativas, problemas de custo e sentimentos. As formas de curva foram categorizadas de acordo com suas tendências para melhorando, deteriorando ou estável. Os motivos também foram classificados como pragmáticos e hedônicos baseados no modelo de experiência do usuário de Hassenzahl (2003). A Tabela 7 mostra os motivos que mudaram o relacionamento com a pulseira e o aplicativo ao longo do tempo. Conforme mencionado estas descrições foram sumariamente classificadas de acordo com o modelo de Hassenzahl (2003). É possível notar nas descrições, que a experiência se altera de acordo com o tempo de uso e/ou interação. Razões pragmáticas e hedônicas mais positivas foram associadas à elevação das curvas e razões pragmáticas e hedônicas negativas foram associadas com o decréscimo das curvas. Isso sugere que a melhoria das curvas está relacionada a uma adequada UX geral. No

que diz respeito às descrições/razões foram listadas dezessete razões, das quais doze (12) foram negativas e sete (5) listadas como positivas (Tabela 7).

Tabela 7: Descrições Motivos

Descrições		Classificação	
Usuário 1	A principio é fácil de usar	Pragmático Usabilidade	Negativo
	O gráfico é legal	Atratividade Beleza	Positivo
	Aa pulseira não incomoda.	Pragmático	Positivo
	Depois de algum tempo incomoda por exemplo: dormir de barriga para baixo	Pragmatico	Negativo
	Não é fácil identificar quando a pulseira funciona ou quando a bateria acaba	Pragmático Usabilidade	Negativo
Usuário 2	Empolgado	Estimulação	Positivo
	Frustrado com a falta de alguns recursos	Expectativa de uso	Negativo
	Animado em utilizar no computador (ler dados)	Estimulação	Negativo
	Chateado não importa dados	Pragmático Utilidade	Negativo
	Utilizei como notificador	Pragmático Usabilidade	Positivo
	Aprendi o LifeLog	Pragmático Eficiência	Positivo
	Incomoda	Pragmatico	Negativo
Perdeu o interesse	Identificação	Negativo	
Usuário 3	Não fez observações		
Usuário 4	Difícil com o celular, acredita que seja por serem marcas distintas	Pragmático Usabilidade	Negativo
	Complicado pois não sabe se os dados estão sendo sincronizados	Pragmático Usabilidade	Negativo
	A funcionalidade de monitoramento de batimento cardiacos não funciona	Pragmático Usabilidade	Negativo
	Não obteve feedbacks que gostaria	Pragmático Utilidade	Negativo
Usuário 5	Não fez observações		

4.4. Resultados UX Curve

Em suma, o método UX Curve foi razoavelmente claro para o grupo de usuários. Todos os participantes conseguiram desenhar as curvas, embora dois deles tenham achado difícil em princípio, sendo necessário incentivo no início. Com relação aos picos relatados, estes foram no total vinte e dois (22), dos quais onze (11) foram positivos, três (3) picos neutros e oito (8) negativos (Tabela 8).

Tabela 8: Síntese Picos Coletados

Ocorrências Picos			
Positivo	Neutro	Negativo	Total
11	3	8	22

As curvas coletadas também foram classificadas como melhorada, estável ou deteriorada. Destas, três delas foram *deteriorando*, mostrando que as expectativas diminuíram ao longo do tempo. Duas curvas foram catalogadas como *melhorando*, não houveram curvas *neutras*.

Tabela 9: Resultado Classificações

Análise das curvas	
Usuário 1	Melhorada
Usuário 2	Deteriorada
Usuário 3	Deteriorada
Usuário 4	Melhorada

Além das análises supracitadas, foi questionado aos usuários se eles acreditam que o método conseguiu avaliar suas experiências ao longo do tempo. Alguns usuários mencionaram que não foi possível descrever o tempo dos eventos exatamente, por exemplo: "foi difícil dividir os eventos na linha do tempo, não lembrei quando algo aconteceu".

Eles também usaram diferentes maneiras de anotar as curvas. Dois dos usuários marcaram motivos positivos que melhoraram sua experiência no pico da curva, após isto, começou a deteriorar-se; enquanto outro usuário marcou na linha central mas citou um motivo negativo. No entanto, a curva UX não é uma representação exata de experiências no período de uso, mas uma descrição aproximada de eventos que são significativos para eles.

4.5. Discussão e Conclusão UX Curve

Os objetivos do teste com UX curve inicialmente eram estimar sua eficácia na avaliação da UX ao longo do tempo.

Os resultados sugerem que o método UX Curve é útil para avaliar a qualidade da experiência do usuário a longo prazo, uma vez que gerou dados qualitativos que descrevem a qualidade da experiência do usuário ao longo do tempo. Todavia, usuários mencionaram dificuldades para lembrar de acontecimentos ocorridos, e relatar a experiência sobre eles. Além dos dados qualitativos, o UX Curve também forneceu dados quantitativos sobre as tendências da experiência do usuário ao longo do tempo. Entretanto, os dados quantitativos são a quantidade de picos esboçados nas curvas; além dos picos, as oscilações das curvas podem revelar impressões significativas a respeito da experiência com o produto/serviço, e estas não são contempladas no UX Curve.

Uma situação comumente citada pelos usuários é o fato de não saber o que estão avaliando. Embora o UX curve seja capaz de avaliar sumariamente a UX ao longo do tempo, quando se trata de avaliar um cenário de IoT ainda é preciso ajustes para este contexto.

5. Conclusão

O estudo de caso foi conduzido realizando a aplicação dos três métodos supracitados à experiência de cinco usuários voluntários que fizeram uso da Sony SmartBand SWR10¹ um dispositivo vestível, que monitora atividades físicas ou sono. Não foram especificados passos a serem seguidos com relação ao uso, apenas que fizessem uso da conexão entre a pulseira e o aplicativo.

Os resultados do AttrakDiff apresentam resultados positivos, no entanto os usuários afirmaram não saber de qual parte do sistema (pulseira/aplicativo) se tratavam as questões, afirmaram não saber associar as partes. Em cenários de IoT comumente temos dois ou mais dispositivos que agem em conjunto, para prover um determinado serviço. Esta dificuldade do método pode interferir de forma negativa na avaliação da UX neste tipo de cenário, que tem como uma de suas fundamentais propriedades a distribuição.

O Emocards trouxe igualmente resultados positivos para o aplicativo e o dispositivo. A avaliação sem a necessidade de escrita foi um ponto ressaltado por diversos usuários, mas também foi fonte de discussões sobre se as figuras apresentadas são capazes de expressar as emoções sentidas. A junção necessária do dispositivo e aplicativo durante o teste, mostra falha no método que não se mostrou eficaz no que diz respeito a intenção futura de avaliar cenários IoT.

O UX Curve, por sua vez, que se propõe a avaliar como a experiência mudou ao longo do tempo, trouxe resultados que aparentam estar mais próximos à experiência sentida pelos usuários. Este mostrou dados negativos e neutros de acordo com sua metodologia. Por exemplo, grande parte dos usuários afirmou não ter interesse em usar mais o produto, nem ter intenção de adquirir. Contudo, o método apontou algumas questões como, a dificuldade de lembrar dos acontecimentos ocorridos mesmo quando estes eram recentes. Uma outra questão é com relação aos dados avaliados, que consideram apenas os picos de alteração. Este tipo de avaliação pode desconsiderar impressões significativas para a avaliação. Por exemplo, no desenho de algumas curvas, os usuários afirmaram ter colocações durante a trajetória da curva, e não somente nos picos.

Contudo, é importante ressaltar que o estudo foi exploratório, com uma pequena amostra de participantes. O tempo de uso do dispositivo variou, o que para um método de avaliação de UX ao longo do tempo pode afetar os resultados finais. É necessário inicialmente a compreensão dos cenários de IoT e suas particularidades. Em seguida a aplicação dos métodos nestes cenários poderá trazer respostas sobre a avaliação da UX nestes cenários.

Referências

- Desmet, Pieter, Kees Overbeeke, and Stefan Tax. "Designing products with added emotional value: Development and application of an approach for research through design." *The design journal* 4.1 (2001): 32-47.
- Desmet, Pieter. "Measuring emotion: Development and application of an instrument to measure emotional responses to products." *Funology*. Springer Netherlands, 2003. 111-123.
- Hassenzahl, M. (2003). The thing and I: Understanding the relationship between user and product. In M. A. Blyth, A. F. Monk, K. Overbeeke, & P. C. Wright (Eds.), *Funology: From usability to enjoyment*, 1-12 (chap. 3). Kluwer Academic Publishers.
- Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human- Computer Interaction*, 19, 319-349.
- Holbrook, m.b. (1985). 'Emotion in the consumption experience: Toward a new model of the human consumer'. In Peterson, Hoyer and Wilson (Eds), *The role of affect in consumer behavior: Emerging theories and applications*. Lexington, MA: Heath.
- Mendoza, V. and Novick, D.G. 2005. Usability over time. *Proceedings of the 23rd annual international conference on Design of communication: documenting & designing for pervasive information* (New York, NY, USA, 2005), 151–158.
- Karapanos, E., Zimmerman, J., Forlizzi, J. and Martens, J.- B. 2009. User experience over time: an initial framework. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2009), 729– 738.
- Kujala, S., Roto, V., Vaananen-Vainio-Mattila, K., Karapanos, E., Sinnela, A.: *UX Curve: A Method for Evaluating Long-Term User Experience*. *Interacting with Computers* (2011).
- Prentice, Deborah A. "Psychological correspondence of possessions, attitudes, and values." *Journal of Personality and Social Psychology* 53.6 (1987): 993.
- Reynolds, T.J. and Gutman, J. (1988). 'Laddering theory, Method, Analysis, and Interpretation'. *Journal of Advertizing Research*, 28, 11-31.
- Russell, j.a. (1980). 'A circumplex model of affect'. *Journal of Personality and Social Psychology*, 39, 1161-1178.
- Schlosberg, h. (1952). 'The description of facial expression in terms of two dimensions'. *Journal of Experimental Psychology*, 44, 229-237.