User Experience Testing: A Case Study for Mobile Devices

Lim Tek Yong
Faculty of Computing and Informatics
Multimedia University
Cyberjaya, Malaysia
tylim@mmu.edu.my

ABSTRACT: *There are plenty of user experience evaluation methods available in the literature. However, most user experience evaluation methods measured different dimensions of user experience. Therefore, Norman's emotional design is proposed as a common conceptual framework for user experience. User experience evaluation method should measure emotional response related to visceral, behavioural and reflective. A case study of user experience testing was carried out and result revealed that the selected mobile device scored high on usable, medium on valuable and low on likeable. Thus, a user experience testing is designed to provide holistic view of a product.*

Keywords: User Experience Testing, Emotional Response, Evaluation Method, Mobile Devices

Received: 3 May 2013, Revised 7 June 2013, Accepted 11 June 2013

1. Introduction

Evaluation plays an important role in product development lifecycle [1]. Evaluation is referring as an activity to assess the value of products with respect to a specific benchmark. It can inform developers whether the product meet or fail certain standards required by public and industry. Academic and industry professionals are constantly searching for evaluation methods that can provide an accurate product assessment [2, 3]. This assessment can be used to gauge the product success in the real market and attract potential customers.

In the human computer interaction domain, [4] identified five generations of evaluation based on a product evolution, particularly the computer. The generations of evaluation were (1) product reliability, (2) product performance, (3) user performance, (4) usability and (5) user experience. In the 1940s, the first generation evaluation was to ensure that a product would perform without failure. Once the product became reliable in the 1950s, the second generation evaluation emphasised on the speed that a product took to process huge data. Then, time-sharing product became popular in the late 1960s and the third generation evaluation began to assess user performance such as task completion rate and task completion time. Later in 1980s, the number of novice users using the product was increased tremendously and the fourth generation evaluation included additional measures such as learnability and others. From 2000 till present, the product became ubiquitous and the focus shifted to user experience. Thus, the current generation evaluation recognised the important of pleasure in product design [5].

2. User Experience

According to the International Organisation for Standard definition, user experience is defined as "*a person's perceptions and responses that result from the use or anticipated use of a product, system or service*" [6]. In other words, it studies on how a customer feels about a particular product that may change over time (including before interaction, during interaction, and after interaction). The customer may experience different states such as positive, negative or neutral. This is also known as emotional responses [7]. Thus, an understanding on various design characteristics of a product that evoke customer's emotional responses required more rigorous research on developing new user experience evaluation.

A study conducted by [8] revealed that 96 user experience evaluation methods were found in the literature. These evaluation methods included 3E (Expressing Experiences and Emotions), aesthetics scale, emocards, Experience Sampling Method (ESM), Geneva Emotion Wheel, Kansei engineering software, Self Assessment Manikin (SAM), and others. Every user experience evaluation method has its own advantages and disadvantages. Besides that,each evaluation method measured different dimensions of user experience. The following section highlights selected user experience evaluation methods found in the literature.

3. User experience Evaluation Methods

Due to the limited space, only four user experience evaluation methods were reviewed, namely SAM, 3E, psychophysiological and ESM.

3.1 SAM

A portion of this submission has been published in the proceedings of 2013 Third International Conference on Innovative Computing Technology.

SAM is a non-verbal method designed by [9]. This method uses a collection of doll images that measures three different emotion scales, namely pleasure, arousal and dominance [11]. This method is easy to use, require only simple equipment (e.g. paper and pencil), and result can be presented in quantitative format. [10] conducted two field experiments using SAM and found that some participants may not interpret the scale correctly. Thus, an actual understanding of product features evoke customer emotional response can be acquired by complement SAM with additional data collection.

3.2 3E

3E is a self-report diary method to express customer emotion [12]. During the evaluation, participants are given a basic template. The template contains a stick character with blank face,speech bubble and thought bubble. Participants can state their emotional response by writing text message and sketching facial gestures. This method provides rich data on emotional response. The disadvantages of this method are researchers need to wait for a period of time until participants returned the diary, researchers would face difficulty in interpreting the diary content, and some participants do not like to draw [10].

3.3 Psychophysiological

Psychophysiological is a laboratory method that combines both subjective and objective data [13]. During the evaluation, sensors are attached to participants to record physiological responses such as heart beat, skin perspiration, and facial muscles. After the evaluation, participants are asked to report their experience and rate their subjective responses in term of boredom, challenging, frustration and fun. But this method requires special equipment (e.g. video splitter, galvanic skin response sensor, and electrocardiography sensor) and expertise knowledge on physiology.

3.4 ESM

ESM is a systematic self-report method that gathers participants' emotional response as they experience in real time [14]. A mobile ESM system, called MyExperience, is capable of detecting more than 140 events (e.g. communication, device usage, user context, location, and environmental sensors). Participants are prompted with questions on a mobile device and they have to answer these questions [15]. The advantage of this method is that activity log can pinpoint the exact design feature evokes the participant's emotional response. In other words, memory effect is minimised. However, both product to be evaluated and data collection tool are required to be running on same platform. Furthermore, participants may feel the ESM system is disturbing in certain situations (e.g. driving).

Most user experience evaluation methods were originated from academic because many industry professionals would not disclose their company methods to the public [8]. Each evaluation method can be categorised into the following attributes: source of the method, type of collected data, type of product being investigated, product development phases, type of participants, evaluation venue, duration of experience and special requirements. Most methods can be used to elicit information from a single person, to evaluate fully functional product/prototype such as mobile app, and conducted without using any special hardware. Half of the methods allowed participants to express their experience freely. Some methods required special equipment, and some data analysis can be quite challenging [8].

However, product developers favoured a practical user experience evaluation method [16]. In developer's point of view, the user experience evaluation method should be easy to deploy and require only low expertise level. It could be used in different stages of product development lifecycle such as concept ideas, prototypes and products. It should be able to cater different target customer groups and different product types. Therefore, there is a need to find user experience evaluation that provides holistic view for a product [4].

Although there was a clear definition on user experience, but product developers still have not agreed on standard measurement for user experience [16]. There was no consensus between researchers on criteria for assessing user experience because each evaluation method measured different emotional responses. Some studies may measure the same emotional response but using different words [17]. So, a common conceptual framework for user experience is useful to address the above issues.

4. Conceptual Framework for User Experience

To date, there are numerous conceptual frameworks for user experience being proposed by researchers from various research backgrounds [18, 19, 20, 21, 22, 23 and 24]. Most frameworks classified emotional response into three different levels [25]. For comparison purpose, Norman's terminology is used as a point of reference, namely visceral level, behavioural level and reflective level (see Table 1).

Authors	Emotional Responses Level		
[18]	Visceral	Behavioural	Reflective
[19]	Intrinsic attractiveness	Semantic attractiveness	Symbolic attractiveness
[20]	Response to form	Response to function	Response to meaning
[21]	Sensory/ aesthetic response	Cognitive/behavioural response	Personal/ symbolic response
[22]	Objects	Agents	Events
[23]	Visual forms	Purposefulness and functionality	Full the need to belong and for self esteem
[24]	Sensory	Interaction and behaviour	Personality

Table 1. Conceptual Framework for User Experience (adapted from [24] and [25])

Visceral level is referring to emotional response based on customer's physiological senses. This level involves human sensory such as see, hear, touch, taste and smell the product. Normally, human would judge a product solely based on its aesthetic appearance and physical characteristics (e.g. shape, colour, texture and others). For instance, customers may like the slim shape of a mobile device. It is also known as a response to product form.

Behavioural level is referring to emotional response based on customer's interaction. This would require human to perform specific tasks using a product in order to achieve certain goals. So, human would assess based on product's interaction technique and their own performance (e.g. effectiveness, efficiency, learnability and others). For example, customers may enjoy the simple steps to take a picture using the mobile device. It is also called a response to product usability.

Reflective level is referring to emotional response based on customer's reflective thought. In this level, human may relate to their past experiences and other products. Human may also associate a product with their personality, financial, lifestyle, culture and ideology. Thus, human would evaluate based on the product identity (e.g. brand, price, trend and others). As an illustration, customers may feel proud using the latest model of branded mobile device. It is also known as a response to product meaning. Product developers should design an ideal product that fulfils all three levels of emotional response. In plain English, an attractive, usable and personalised product could lead decision to product ownership. In real market place, some customers are willing to pay more for this product. Many researchers identified mobile device as an interactive product to be assessed in user experience evaluation [10, 26, 27 and 28]. In user experience research, a product type should not be limited only for work related usage but the product also being able to support leisure activities [17]. For example, mobile devices can be used by customers for leisure (e.g. playing mobile game), work (e.g. making emergency call) and both.

In this paper, mobile device is defined as a handheld size computer with an operating system and capable to run various applications. This included smart phone, tablet, personal digital assistant and others. Few previous studies proposed different user experience evaluation methods to examine how human express their emotion for mobile devices in these three levels [27, 29 and 30].

5. User Experience Evaluation Methods for Mobile Devices

Two main approaches were employed by previous researchers to study emotional response levels, namely exploring each level separately and exploring all levels at one time.

5.1 Exploring Each Emotional Response Level Separately

[27] suggested that each level of emotional response should be gauged separately using different user experience evaluation methods. For the visceral level, mobile developers can employ a think aloud protocol. This technique involves observing customers experience a mobile device for the first time and at the same time, customers are require to verbalise their inner thought. Mobile developers are recommended to employ ESM for the behavioural level due to contextual factors. The researchers argued that interaction between customer and mobile device can elicit different emotional responses in different context. Interviews and questionnaires are suggested for the reflective level. Mobile developers can probe customer's experience using verbal and non-verbal measurement tools. Besides that, the researchers also proposed that interviews and questionnaires can be used for pre-experience level. As a result, mobile developers have to carry out more evaluation studies in order to understand each emotional response level. Indirectly, this implies more time and more people are needed.

5.2 Exploring All Emotional Response Levels at One Time

[29] employed a methodological triangulation that combined photo diary, interview and questionnaire methods. In this study, twelve participants were asked to take picture of interactive products. After that, the participants were interviewed about their demographic data and general impressions on the interactive products (e.g. why they like a particular product). Then, the participants were asked to share their experiences and rate the selected products using semantic differential scale. The result revealed that these emotional response levels were clearly unique from one another. Visceral level was tightly associated with visual quality (e.g. colour), behavioural level was linked with interaction quality (e.g. interfaces) and reflective level was highly related to functional quality (e.g. making call).

On the other hand, [30] exploited online reviews to identify different levels of emotional response for iPhone. This method is also considered as indirect observation where online reviewer performed his/her activity in natural setting. The output of reviewer activity was made available online. Based on ten different online reviews, the researchers identified emotional response used (e.g. shiny) and associated design feature (e.g. colour). Then, the researchers mapped their findings into different levels of emotional response. The study revealed that online reviewers also expressed positive and negative emotional responses in the review. However, the availability of online review is solely depending on reviewer effort.

Although there is a need to find alternative evaluation methods to uncover user experience, but the existing usability evaluation methods can be tweak to include emotion elicitation. The reason is that usability still needed in product development and it should not be discarded from evaluation. So, usability testing is selected as an example. Usability testing is an evaluation method to test whether a product is usable for a specific user group to perform certain tasks. In this paper, user experience testing is an adapted method from traditional usability testing. The new user experience testing is not only for detecting usability bugs but also uncovering "*negative user experience*".

6. User Experience Testing: A Case Study

A case study was undertaken to show that the user experience testing can be used to identify negative user experiences and usability problems.

6.1 Mobile Device

There are many types of mobile device available in the market. Samsung Galaxy W (i8150) has been selected as a case study for user experience testing. This mobile device is considered as a smart phone because it running the Android operating system. It was chosen based on state-of-the-art smart phone functions such as touch screen with resolution 480 x 800 pixels, 5 megapixel camera and music player. This mobile device was available in market at the end of 2011.

6.2 Test Participants

An initial screening was carried out in Multimedia University campus and only potential participants that with no experience with the selected mobile device were recruited. Five test participants joined in this testing. Two test participants were male and the remaining three were female. They were from 19 to 22 years old and their average age was 20.2 years. All test participants were undergraduate students and lived in Cyberjaya, Selangor, Malaysia. They had used mobile phones for more than three years. Three were iPhone users, one was Blackberry user and one was Sony user. A token of appreciation was given to each test participant for their time.

6.3 Tasks

In this user experience testing, each test participant was asked to perform four tasks using the selected mobile device. The tasks given were as follow:

Task 1: Give your first impression of this mobile device. Feel free to explore this device as you normally would, but please don't unlock the device just yet.

Task 2: Call mobile service operator to find out the unbilled amount for this device number. The number that you need to call is 016 221 1800.

Task 3: Take a creative photo of the lab. You should inform moderator on your final photo if you took more than one.

Task 4: Decide whether you want to own this mobile device or not. A detailed technical specification of this mobile device is provided for your reference.

Task 1 is used to elicit emotional responses at visceral level. Test participants may express their liking or disliking based on minimum of information [33]. For eliciting emotional responses at behavioral level, Task 2 and 3 are used. Task 2 is considered as work-related and mobile users used about 12.3 voice calls per day [35]. Task 3 is focused on leisure aspect and also one of the popular activities among mobile users [34]. Task 4 is used to elicit emotional response in reflective level. Test participants may value the mobile device by comparing the retail price with the specification [36].

6.4 Procedure

The user experience testing was conducted in a controlled lab setting at Multimedia University. Firstly, a moderator welcomed test participants and a brief description on the goal of user experience testing was provided. Then, test participants were requested to read and sign a consent form. They were allowed to leave from the testing at any time, their identity and personal data collected would be kept confidential. Subsequently, the testing began with moderator introduced a mobile device and explained the think aloud protocol. Test participants were asked to comment on what they were looking at, what were their inner thought, what were they doing, and how did they feel about the given tasks. Test participants were also reminded that the testing was about mobile device and not them. They should feel free to express anything, whether positive or negative experience. Each test participant was given a set of tasks to complete with the mobile device. Whenever test participants stopped talking, moderator would prompt them to "*keep on talking*". During the testing, a video camera was used to record test participant's interactions and gestures. A voice recorder was also used as backup to record test participant's think-aloud commentary. Participants were given maximum five minutes to complete each task. Lastly, test participants were interviewed by moderator.

6.5 Measurement

A binary like-dislike scale was used as subjective measurement for Task 1 and test participants were required to justify their reason with reference to any design characteristic. For task 2 and task 3, two objective measurements were used, namely, successful task completion and time on task. A task completion is considered successful when test participants indicated that they completed the task. Time on task is measured based on the amount of time spent for a participant to complete the task. A binary cheap-expensive scale was used for Task 4 and test participants were asked to give reason related to any design characteristic.

Severity rating was used to indicate the quality of user experiences provided by the mobile device such as likeable, usable and valuable. The quality is measured based on three point scales namely, low, medium and high. In this testing, high means any issue that annoys only one or two persons. Medium means any issue that nearly half of the sample experienced the same annoyance. Low means any issue that most or all facing the annoyance.

6.6 Result

All test participants had a poor first impression of the selected mobile device in Task 1. Design characteristics that failed to attract test participants were device shape, colour and material. Three test participants disliked device shape and expressed it as old fashion, ugly design and unattractive. One test participant disliked plain black colour and make the device looked boring. Another test participant indicated that the device quality was low after touched plastic material. In short, the mobile device was rated as low on visceral level.

Task 2 was considered simple task and all test participants successfully completed the task less than two minutes. Although they were not familiar with the mobile device, they managed to find 'phone' icon on bottom of the screen. They found the task was easy and the call quality was good. All test participants successfully took a creative photo for Task 3. Three test participants took more than one photo and resulted in longer time completion. Two test participants informed that photo was blurring due to hand movement. They realized this issue after reviewed the photo taken. In addition, all test participants expressed that there was a short delay when starting up the camera app. Both problems were considered as minor. In sum, the mobile device was rated as high in behavioral level.

In Task 4, three test participants did not intent to own the mobile device. Their reasons were concerned on small size of internal storage, expensive price compare to low build quality, and willing to fork out few hundreds more to get a better smart phone with better specifications. The remaining two test participants were satisfied with the current specifications and wish to own the mobile device. In other words, the mobile device rated as medium on reflective level.

6.7 Suggestion for improvement

Based on the findings, the selected mobile device achieved good usability. Minor problems found in photo taking can be resolved by implementing image stabilizer and immediate feedback such as loading status. However, two major improvements are needed for the selected mobile device, namely visceral and reflective levels. For the visceral level, developers and designers can look into current shape, colour and material used. They may explore other visual designs and research on current trend mobile device. Otherwise, they may also consider physical customization where potential customers can choose their preferred shape, colour and material. On the reflective level, engineering and marketing team may need to revise or repackage on internal storage size, build quality and additional features. The above recommendation is more suitable if the company only launch a single model. For company like Samsung, promoting many models may resolve certain issue highlighted in reflective level.

7. Discussion

The goal of proposed user experience testing is to discover how real users experience a mobile device. At the same time, mobile developers can identify potential usability bugs and negative user experience. Mobile developers are able to use the evaluation findings to improve their specific design weakness. It can provide a holistic assessment and can merge with other objective measurements (e.g. heart rate). The following are steps for conducting user experience testing.

First, mobile developers need to identify potential users groups and recruit these users as test participants. A minimum five test participants should be sufficient to uncover the major problems in a mobile device. Test participants are asked to sign an informed consent form before taking part in user experience testing. The user experience testing should begin with visceral level, follow by behavioural level and end with reflective level. During the testing, tests participants are encouraged to think aloud. Both objective and subjective measurements can be used in this user experience testing (see Table 2).

Task Examples	Objective Measurement	SubjectiveMeasurement	Design Characteristics
Give an initial impression	-	Attraction effect (e.g. like/dislike)	Physical Outlook (e.g. Colour)
Perform certain tasks	User performance effect (e.g. time taken to complete)	Psychological effect (e.g. enjoyment/frustration)	Interaction Technique (e.g. Navigation)
Decide product ownership	-	Value effect (e.g. cheap/expensive)	Identity (e.g. price)

Table 2. Proposed User Experience Testing

For the visceral level, test participants are given short time duration to look, touch and feel a mobile device (see Figure 1). Then, test participants are asked to give an initial impression for selected mobile device. Attraction effect (e.g. like or dislike), can be used as subjective measurement. Then, test participants are probed to indicate their reason by referring to mobile design characteristics, especially physical outlook (e.g. colour).

Figure 1. Visceral Level

In the behavioural level, test participants are asked to perform few common tasks (e.g. make a call) using the mobile device (see Figure 2). The task difficulty can ranged from easy to hard. So, usability aspect, such as user performance, can be objectively measured. This included time taken to complete, task completion rate and others. At the end of each task, test participants are asked to indicate their psychological effect (e.g. their enjoyment or frustration level). Test participants need to justify their emotional response by relating to the interaction technique (e.g. navigation using swipe gesture) provided in mobile device.

Test participants are requested to decide whether they wish to own the mobile device in the reflective level. Additional information (e.g. technical specification and promotional leaflet) can be provided to test participants for consideration (see Figure 3). Test participants can point out their value for mobile device (e.g. cheap or expensive). They can give their justification by associating with the mobile identify (e.g. price, brand). It is recommended that test participants should classify each subjective response either a positive or negative experience. This is because “*cheap*” can mean good value and reasonable. For others, “*cheap*” can also refer to substandard and inferior quality.

Test participants can complete a user experience testing within thirty minutes to one hour. Sometimes, the testing may take longer due to unforeseen circumstances such as equipment failure, participant engrossment, participants who arrive late and other factors. So, mobile developers should allocate sufficient time between testing sessions. In a long user experience testing,

Figure 2. Behavioural Level

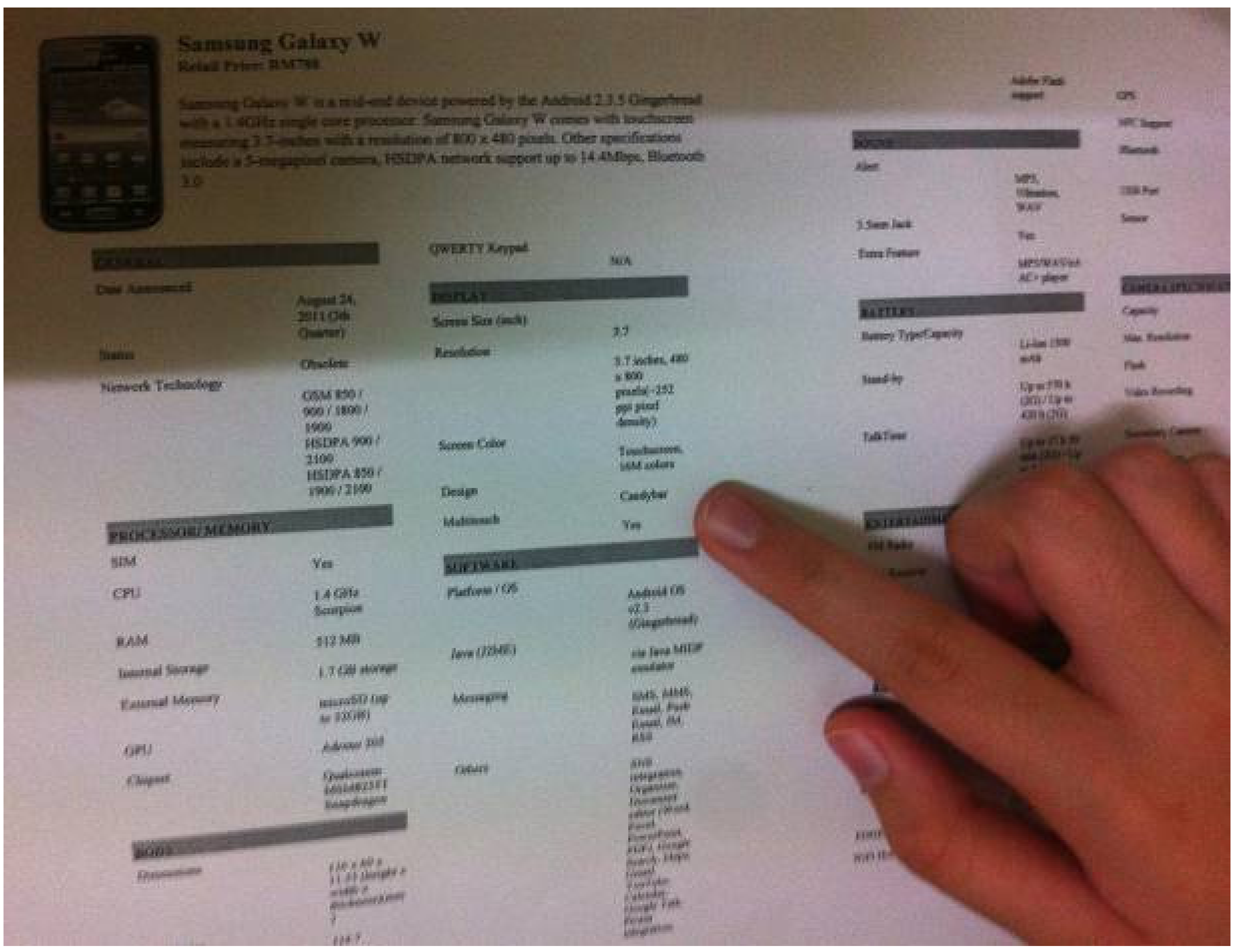

Figure 3. Reflective Level

test participants may be tired or bored and this eventually affects their emotions. So, mobile developers should conduct a pilot study to ensure the user experience testing is feasible in term of time, cost, instruction, equipment and others.

The user experience testing can be carried out in a conventional usability lab or a controlled environment setting. This means test participants and mobile developers are in the same physical location. A camcorder is utilised to record test participant's interaction during the testing and additional web camera can be used to record test participant's facial expression. In addition, test participants' subjective responses can be recorded on a piece of paper or typed in electronic form. Although paper-and pencil method is cost effective, mobile developers may face difficulty in reading poor handwriting. So, electronic form is preferable. Besides, misspelled words can be flagged and corrected during data entry. Electronic form can also eliminate the data entry from paper and data analysis can begin immediately since the data is already stored in database.

Mobile developers need to examine both qualitative and quantitative data at the end of user experience testing. They can use spreadsheet software to perform a simple quantitative analysis. For example, result for time taken to complete can be reported in the form of averages and percentages. In addition, various graphical representations can be generated using the spreadsheet software. On the other hand, Nvivo software can be used to support analysis of qualitative data especially the video clips and open ended responses. This software allows mobile developers to label, organise and manipulate the subjective emotional responses more effectively. For instance, mobile developers can choose to identify recurring patterns or look for critical incidents. A list of usability problems and negative user experiences can be compiled. A electronic report should present the findings based on three levels of emotional responses. Screenshot or short video clips can be used to illustrate specific negative user experience faced by test participants. A recommended solution should be provided for each negative user experience. A severity rating can also be included to indicate the level of likeable (visceral), usable (behavioral) and valuable (reflective) for a mobile device.

However, the proposed user experience testing may not suitable for mobile developers who seek for new design inspiration. The user experience testing result only indicate that test participants may have negative experience associated to a particular design characteristic. In addition, reflective level is more complex than it seem. Mobile developers may need to consider other human value issues like ethics, security, social, cultural, personal development, and others. The following sub-sections discussed on other challenges faced in user experience evaluation method.

7.1 Test Participants

Test participants in user experience evaluation may have different viewpoint. They are not product designer as they had limited knowledge in design. Test participants may only evaluate the product as a whole and give a general emotional response. Thus, mobile developer may encounter emotional response such as *"I wish that it comes in different colour like pink"*. For a mobile developer, "*pink*" colour can be in different hue, tint, and shades. So, the mobile developer may have to decide whether the test participant was referring to bright pink, dark pink, light pink, ultra pink, or shocking pink.

7.2 External factors

Company usually imposes logistics constraints for each product development [31]. Product development team is given limited resources such as time, money and people. Mobile developers want to collect feedback from participants as soon as possible. Then, mobile developers can use test participant's feedback as input for improving their product. If an evaluation method consumes long duration, then the process of data gathering may slow down the product development lifecycle. Furthermore, the evaluation result may not be applicable as new trend or technology enters the market.

7.3 Context dependent

Emotional response collected for certain product may not be applicable in other contexts although the product remains the same. This may due to different culture, different location, different time, different age, different personality and others. For example, when mobile users misplaced their mobile device in their home, they would be happy when they heard the sound of ringtone. However, they would respond differently to the same ringtone in a different location. They would be embarrassed to hear the same ringtone when they were in a meeting room. Thus, mobile developers need to consider and explore other possible contexts.

7.4 Long term usage

As mobile users used mobile device for a period of time, there would be other emotional responses associated to the mobile device, such as panic, strangeness, being cool, irrational behaviour, thrill and anxiety [32]. When mobile users are separated from their mobile device or battery is dead, mobile users become panic easily. The possible reason is because mobile device contains important data about the mobile users' life. This may include calendar, contact list, messages, banking, emails, photos, documents, and web history. In this context, mobile developers can explore and design more security features in mobile devices.

8. Conclusion

This paper highlighted various user experience evaluation methods to measure different levels of emotional response. A practical user experience evaluation method can help product developers to assess their prototype or product effectively. Future research should look into possible inspection method for user experience especially when user experience specialists can be invited to evaluate the product [4].

9. Acknowledgment

This research was supported by the Ministry of Higher Education Malaysia (Grant No. FRGS/2/2010/SG/MMU/03/4).

References

[1] Moreno, M., Seffah, R., Capilla, Sanchez-Segura, M. (2013). HCI Practices for Building Usable Software, *Computer*, 46 (4) 100-102.

[2] Greenberg, S., Buxton, B. (2008). Usability Evaluation Considered Harmful (Some of the Time), *In*: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p. 111-120.

[3] Roto, V., Vaananen-Vainio-Mattila, K., Law, E., Vermeeren, A. (2009). User Experience Evaluation Methods in Product Development (UXEM'09), *In*: Proceeding of the INTERACT, Part II, LNCS 5727, p. 981–982.

[4] MacDonald, C. M., Atwood, M. E. (2013). Changing Perspectives on Evaluation in HCI: Past, Present and Future, *In*: Proceedings of the CHI '13 Extended Abstracts on Human Factors in Computing Systems, p. 1969-1978.

[5] Allam, A. H., Hussin, A. R. C., Dahlan, H. M. (2013). User Experience: Challenges and Opportunities, *Journal of Information Systems Research and Innovation*, 3, February, p. 28-36.

[6] International Organization for Standardization, ISO 9241-210. (2010). Ergonomics of human-system interaction - part 210: human-centred design for interactive systems. Switzerland.

[7] Agarwal, A., Meyer, A. (2009). Beyond Usability: Evaluating Emotional Response as an Integral Part of the User Experience. *In*: Proceedings of CHI '09 Extended Abstracts on Human Factors in Computing Systems, p. 2919-2930.

[8] Vermeeren, A. P. O. S., Law, E. L., Roto, V., Obrist, M., Hoonhout, J., Vaananen-Vainio-Mattila, K. (2010). User Experience Evaluation Methods: Current State and Development Needs, *In*: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, p. 521-530.

[9] Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: computer applications, Technology in Mental Health Care Delivery Systems, p. 119-139.

[10] Isomursu, M., Tahti, M., Vainamo, S., Kuutti, K. (2007). Experimental evaluation of five methods for collecting emotions in field settings with mobile applications, *International Journal of Human-Computer Studies*, 65 (4) 404–418.

[11] The PXLab Self-Assessment-Manikin Scales, Retrieved from http://irtel.uni-mannheim.de/pxlab/demos/index_SAM.html.

[12] Tahti, M., Niemelä, N. (2006). 3e – expressing emotions and experiences, In Proceedings of the WP9 Workshop on Innovative Approaches for Evaluating Affective Systems, Retrieved from http://emotion-research.net/projects/humaine/ws/wp9/d9e.pdf.

[13] Mandryk, R. L., Inkpen, K. M., Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour and Information Technology*, 25 (2) 141-158.

[14] Larson, R., Csikszentmihalyi, M. (1983). The experience sampling method. *In*: New Directions for Methodology of Social and Behavioral Science, 15, p. 41-56.

[15] Froehlich, J., Chen, M. Y., Consolvo, S., Harrison, B., Landay, J. A. (2007). MyExperience: A System for In situ Tracing and Capturing of User Feedback on Mobile Phones, *In*: Proceedings of the 5th International Conference on Mobile Systems, Applications and Services, p. 57-70.

[16] Vaananen-Vainio-Mattila, K., Roto, V., Hassenzahl, M. (2008). Towards Practical User Experience Evaluation Methods. *In*: Proceedings of International Workshop on Meaningful Measures: Valid Useful User Experience Measurement, p. 19-22.

[17] Bargas-Avila, J. A., Hornbaek, K. (2011). Old Wine in New Bottles or Novel Challenges? A Critical Analysis of Empirical Studies of User Experience, *In*: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p. 2689-2698.

[18] Norman, D. (2004). A Emotional Design: Why We Love (or Hate) Everyday Things, Basic Books.

[19] Baxter, M. (1995). Product Design: A Practical Guide to Systematic Methods of New Product Development. Chapman & Hall.

[20] Crozier, R. (1994). Manufactured Pleasures: Psychological Response to Design. Manchester University Press, Manchester, UK.

[21] Cupchik, G. C. (1999). Emotion and Industrial Design: Reconciling Means and Feelings. *In*: Proceedings of the 1st International Conference on Deisgn and Emotion, p. 75-82.

[22] Desmet, P. R. (2002). *Designing Emotions, PhD thesis, Delft University of Technology*, the Netherlands.

[23] Lewalski, Z. M. (1988). Product Aesthetics: An Interpretation for Designers. Design & Development Engineering Press.

[24] Gorp, T. V. (2008). Understanding Design for Emotion Models, Retrieved from http://www.affectivedesign.org/archives/199

[25] Wrigley, C. (2013). Design Dialogue: The Visceral Hedonic Rhetoric Framework, *Design Issues*, 29 (2) 82-95.

[26] Mansoor, A., Mahboob, Z. (2011). User Experience Evaluation in Mobile Industry, Master Thesis.

[27] Obrist, M., Meschtscherjakov, A., Tscheligi, M. (2010). User Experience Evaluation in the Mobile Context, Mobile TV: Customizing Content and Experience, p. 195-204.

[28] Nakhimovsky, Y., Eckles, D., Riegelsberger, J. (2009). Mobile user experience research: challenges, methods & tools, *In*: Proceedings of CHI '09 Extended Abstracts on Human Factors in Computing Systems, p. 4795-4798.

[29] Y. Lim, J. Donaldson, H. Jung, B. Kunz, D. Royer, Ramalingam, S., Thirumaran, S., Stolterman, E. (2008). Emotional Experience and Interaction Design, *Affect and Emotion in Human-Computer Interaction*, p. 116 – 129..

[30] Saket, B., Yong, L., Behrang, F. (2012). Toward Emotional Design: An Exploratory Study of IPhone 4, Advances in Affective and Pleasurable Design, p. 480-489.

[31] Crilly, N., Clarkson, P. J. (2006). The influence of consumer research on product aesthetics. *In*: Proceedings of the 9th International Design Conference, p. 689-696.

[32] Vincent, J. (2006). Emotional Attachment and Mobile Phones, Knowledge, *Technology and Policy*, 19 (1) 39-44.

[33] Lindgaard, G., Fernandes, G., Dudek, C., Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good ûrst impression, *Behaviour & Information Technology*, 25 (2) 115–126.

[34] Duggan, M., Rainie, L. (2012). Cell Phone Activities, Pew Research Center's Internet & American Life Project, Retrieved from http://www.pewinternet.org/Reports/2012/Cell-Activities.aspx.

[35] Smith, A. (2011). Americans and text messaging. Pew Research Center's Internet & American Life Project, Retrieved from http://pewinternet.org/Reports/2011/Cell-Phone-Texting-2011.aspx.

[36] Cockton, G. (2006). Designing Worth is Worth Designing. *In*: Proceedings of 4th Nordic conference on Human-computer interaction: changing roles, p. 165-174.