Considering Users' Different Knowledge About Products to Improve a UX Evaluation Method Based on Mental Models

Stefano Filippi[(✉)] and Daniela Barattin

DPIA Department, University of Udine, 33100 Udine, Italy
`{stefano.filippi, daniela.barattin}@uniud.it`

Abstract. User experience (UX) evaluation and design have become important components of the product development process. The UX describes the users' subjective perceptions, responses and emotional reactions while interacting with products or services, considering users' emotions and cognitive activities as the basic elements of the experience. The emotions encompass physiological, affective, behavioral, and cognitive components; the cognitive activities generate and exploit the mental models that govern the human behavior. The literature offers several methods to evaluate the UX; one of them, the irMMs-based method, considers both users' emotions and mental models to evaluate the quality of the UX. Nevertheless, its current release misses the contribution of users who already know the product under evaluation. This research aims at improving the irMMs-based method by considering also users familiar with those products. The expected benefits of this improvement refer to the completeness of the evaluation results and to the definition of relationships between these results and the evaluation activities that allow them to be discovered. All of this can be useful for both researchers and designers who are willing to increase their knowledge about the generation and exploitation of mental models and to select the most suitable evaluation activities to perform time by time depending on the characteristics of the results they are interested in and on the resources available.

Keywords: User experience · Mental models · UX evaluation

1 Introduction

In recent years, user experience (UX) has gained importance in product development because of the increasing number of functionalities (with the related interface complexity), the development of new interaction paradigms, the availability of innovative technologies and devices, etc. The UX describes the users' subjective perceptions, responses and emotional reactions while interacting with products or services, considering users' emotions and cognitive activities as the basic elements of the experience [1, 2]. Emotions are defined as the responses to events deemed relevant to the needs, goals, or concerns of an individual and encompass physiological, affective, behavioral, and cognitive components [3]. Cognitive activities, highly influenced by emotions, generate and exploit the mental models that govern the human behavior. Mental models

are internal representations of the reality aimed at choosing the best user behaviors respect to foreseen behaviors of products or events [4]. User behaviors aim at overcoming specific, critical situations and at achieving experiences as positive and satisfactory as possible [5].

Product development processes need methods and tools for evaluating the UX. The literature already offers several UX evaluation possibilities, from simple questionnaires like the User Experience Questionnaire - UEQ [6] and the IPTV-UX Questionnaire [7] up to more articulated and complete approaches like the Valence method [8] and the irMMs-based method [9], the starting point of this research. The irMMs-based method considers emotions and mental models to evaluate the quality of the UX. IrMMs means interaction-related mental models. These mental models refer to specific situations of interaction involving meanings and emotions as well as users and products' behaviors. The adoption of the irMMs-based method generates lists of positive and negative UX aspects. The former represent unexpected, surprising and interesting characteristics of the interaction; the latter can refer to bad product transparency during interaction, gaps between the actions expected by the user problem solving process and those allowed by the real products, etc.

Indeed, the UX evaluation cannot deserve the time factor. The literature already demonstrated that the subjective evaluation of products and the related subjective meanings assigned to the experiences with them change after an extended period of usage [10]. In their framework named temporary of experience, Karapanos et al. [11] describe how the UX changes over time and highlight three main forces: the increasing familiarity, the functional dependency and the emotional attachment. Sng et al. [12] exploit the UX lifecycle model ContinUE to demonstrate that the 76% of the needs highlighted by the users after a repeated use of a product are completely new respect to those highlighted by users who get in touch with the same product for the first time.

The current release of the irMMs-based method does not involve users who already know the product under evaluation; therefore, it completely misses possible suggestions about the UX quality coming from extended periods of use. This research aims at filling this gap by improving the irMMs-based method considering also users familiar with the products under evaluation. The expected benefits of this improvement refer to the completeness of the evaluation results and to the definition of relationships between these results and the evaluation activities that allow them to be discovered. All of this can be useful for researchers who are willing to increase their knowledge about the generation and exploitation of mental models by users with different knowledge about the products under evaluation, as well as for designers who can understand and select the most suitable evaluation activities to perform time by time, depending on the characteristics of the results they are interested in and on the resources available.

To achieve this goal, the research analyzes the current release of the irMMs-based method to classify the existing evaluation activities and to add the new ones involving users who already know the products. After that, the new release of the irMMs-based method is adopted in the field. The results of the adoption are then compared to those coming from the old release to evaluate the improvements and analyzed to highlight possible relationships with the evaluation activities.

The paper goes on with the background section describing the current release of the irMMs-based method. The activities section describes the improvement of the

irMMs-based method, the first validation of it and the identification of the relationships. The conclusions, together with some perspectives on future work, close the paper.

2 Background: Current Release of the irMMs-Based Method

The irMMs-based method evaluates the UX by exploiting the interaction-related mental models (irMMs). The irMMs consist of lists of users' meanings and emotions along with users' actions and products' reactions/feedback determined by these meanings and emotions [9]. The generation of an irMM comes during the user's attempt to satisfy a specific need in a specific situation of interaction. This process develops through five steps, based on the Norman's model of the seven stages of the action cycle [13]. In the first step, the user perceives and interprets the need. Thanks to this interpretation, in the second step the user recovers existing irMMs from his/her mind and selects those ones that could be suitable to satisfy the need. This selection could be influenced by the presence/consideration of one or more real products. The selected irMMs allow the user to define the goals in the third step. The goals represent intermediate and time-ordered results to achieve; they help in establishing the path to follow to satisfy the need. In the fourth step, the user associates desired meanings and emotions to each goal. These come from the elaboration of the meanings and emotions belonging to the irMMs selected before. Positive meanings and emotions tend to remain as they were; on the contrary, negative ones could be changed into their positive correspondences depending on the number of past experiences that report those negative meanings and emotions. In the fifth step, the user generates the ordered list of actions and the related product reactions/feedback in order to achieve those desired meanings and emotions. This list is generated by elaborating the actions and reactions/feedback present again in the irMMs selected before.

The adoption of the irMMs-based method generates positive and negative UX aspects. This comes by performing tests where real users compare their irMMs, generated before to get in touch with the product as soon as they are told the need to achieve, with the real experience allowed by that product. The irMMs-based method develops through the following four phases.

Phase 1. Input Setting. This phase defines two input. The features of the product to evaluate are the first input; they can be functions, specific procedures to achieve specific goals, physical components, etc. The second input are the users who undergo the tests. Their selection comes by obeying to the rules "the users must not know the product" and "the users must have a specific level of knowledge about the field the product belongs to". Moreover, further rules can apply due to specific characteristics of the evaluations.

Phase 2. Material and Environment Setup. The second phase prepares the material and the environment to perform the tests. The material consists of two documents, one for the users and one for the evaluators. The first document guides the users in generating the irMMs and in performing the interaction with the product. Its first part reports the need to satisfy and the instructions on how to perform each step to generate the irMM, together with examples. The second part shows suggestions on how to

perform the interaction and to compare it with the irMM. This document contains tables and empty spaces to describe the irMM and to comment the comparison. The second document allows the evaluators to collect data in specific tables, divided respect to their type (meanings, emotions, etc.); moreover, this document contains instructions on how to conduct the data analysis. Finally, the test execution requires the setup of a suitable environment reflecting the common use of the product under evaluation.

Phase 3. Test Execution. In the third phase, the users generate their irMMs based on a specific need given by the evaluators and, after that, they compare these irMMs to the real experience allowed by the product. The evaluators collect data in the meantime. Before the tests, the users are divided into two groups. The first group do not get in touch with the product before to know the need to satisfy. This group is called absolute. On the contrary, each user belonging to the second group is asked for interacting freely with the product for a short period. This group is labelled as relative. This label comes from the fact that the product will likely influence the users. Once this free interaction comes to the end, the execution of the tests is the same for every member of the two groups. This execution consists of three moments. In the first moment, each user generates his/her irMM by following the guide document. Then, he/she tries to satisfy the need by using the product (second moment). The user must execute the actions as described in his/her irMM and check if his/her actions and the reactions/feedback of the product come as expected. Of course, problems can occur in the meantime; these problems are addressed as gaps. If a gap shows up, the evaluators suggest the way to overcome the problem in terms of user actions allowed by the product and related reactions/feedback. The user reports the reasons for the gap from his/her point of view and judges if the allowed actions and the product reactions/feedback are better or worse than those of his/her irMM. Once the interaction finishes, the third moment consists in a debriefing where the users are called to reason about their experiences and to express further comments about them. In particular, the users are asked to reconsider the meanings and emotions expressed during the generation of the irMMs in order to highlight any change.

Phase 4. Data Analysis. Some rules lead the analysis of the data collected during the tests. These rules allow generating the positive and negative UX aspects starting from the desired meanings and emotions, the real ones, and the gaps between the expected user actions and product reactions/feedback (those of the irMMs) and the real ones. For example, applying one of the rules, a positive UX aspect is generated starting from the desired meaning "temperature" associated to the goal "washing program for delicate clothes set" and the need "washing delicate clothes with the washing machine". The user gives a positive judgment to the desired meaning because of the reason "although I do not know exactly the most suitable temperature for any type of clothes, the washing machine has predefined programs where all parameters are already set, temperature included". This meaning and the related reason allow generating the positive UX aspect "the washing machine offers predefined programs where all the parameters are already set; the user must only select the right program according to the type of clothes to wash".

  Once generated the UX aspects, they are compared to each other to delete repetitions. They are placed in an ordered list due to the number of occurrences of each UX

aspect, the interaction topics they refer to and their estimated impact. If an UX aspect refers to product characteristics rarely involved in the interaction, its impact will be set to low; on the contrary, if the UX aspect deals with core procedures determining the product cognitive compatibility with the users' problem solving processes, its importance will be higher. The result of the adoption of the irMMs-based method consists of two importance-ordered lists of positive and negative UX aspects.

3 Activities

To achieve the goal of the research, the irMMs-based method (the method, hereafter) must be patched in order to separate the different evaluation activities currently available. This patching makes the method modular and scalable and allows it to be refined by adding the new evaluation activities devoted to the familiar users. We call these activities patching and refining because the former happen at the same level of abstraction while the latter make the abstraction level deeper [14]. At the end, the new release of the method is adopted in the field. This adoption highlights the improvements respect to the old release in terms of quantity and quality of the results and allows identifying the relationships between the UX aspects and the evaluation activities that allowed them to be highlighted.

3.1 Patching the Method

In order to make the different evaluation activities currently available clearly separated, the patching comes by analyzing each phase of the method adoption. Once highlighted, the activities are collected into two separate modules called absolute beginners (AB) and relative beginners (RB) modules, depending on the group of testers involved. This analysis also allows some activities to be revised and optimized. In the following, the peculiarities of these two modules are reported, phase by phase.

Regarding the first phase, concerning the input setting, the AB and RB modules are the same; they share the input in terms of product features and rules to select the users.

The second phase, concerning the material and environment setup, needs more attention. The guide document is the same; nevertheless, it must be diversified between the two modules. The guide document used in the AB module must be made free of references to the product to evaluate to avoid bias. On the contrary, the guide document in the RB module must contain precise information on the characteristics of the product under evaluation to make their attention focusing more on it than on those they could have used in the past. Focusing now on the document for the evaluators, its structure needs some changes in order to integrate the nature of the information (meaning, emotion or gap) with the moment of its generation and the user who finds it, since this is the required information to address the origin of each UX aspect precisely.

In the third phase, where the test execution takes place, the activities in the modules must come as described in the specific guide documents generated in the second phase.

The fourth phase, the data analysis, comes in the same way for both the modules, by filling the documents prepared for the evaluators in the previous phase.

3.2   Refining the Method

The second research activity refines the method by adding the new module devoted to the users familiar with the product. This module is called relative expert (RE) module. This addition makes the method more complete because these users consider the product from a different point of view. For this reason, they could find different UX aspects. What follows describes the peculiarities of the RE module, phase by phase.

In the first phase, the features of the product to evaluate are the same as for the AB and RB modules. On the contrary, the users' characteristics are different; they must know the product and have used it at least for a given period (its duration varies time by time, respect to the product complexity, the required reliability of the evaluation results, etc.). From the point of view of the method structure, the refining requires two more input to introduce, representing the selected modules to adopt time by time and the characteristics of the evaluators, respectively. The new release of the method will be flexible enough to allow the selection of the modules to adopt depending on the specific evaluation. This is the reason for the first new input. Regarding the second new input, although in the case of the AB and RB modules the evaluators can be any, those involved in the RE module must be very skilled and knowledgeable with the product. This is why we need an input to represent the characteristics of the evaluators.

In the second phase, the guide document for the expert users must contain precise references to the product under evaluation in all the activities of the irMM generation (definition of the goals, meanings, emotions, actions and product reactions/feedback). This because the irMMs must focus on the product under evaluation rather than being influenced by past experiences of the users with different products.

In the third phase, the activities of the RE module consist of the same three moments as the AB and RB ones. Thanks to the high level of expertise of the evaluators with the product, the execution of the irMMs can be performed without problems because in case of users running in very specific gaps the evaluators are able to overcome them easily. Finally, during the tests, the evaluators must push the users to stay focused on the specific product features established as input because the high level of knowledge about the product could make the users move to different features.

Finally, the RE module performs the fourth phase by following the same rules as the AB and RB ones by filling its dedicated document and generating the lists of positive and negative UX aspects.

At the end of the refining activities, the new release of the method consists in the three modules, AB, RB and RE. The evaluators can choose any combination of them due to resources available, personal evaluation style, etc. After the selection, the evaluators perform the four phases for each module as described in the previous sections. Since the modules are separated, they can be adopted in parallel, providing that no influences happen among them. At the end, the UX aspects generated by the modules adopted for the evaluation are collected together, compared to each other to delete repetitions and ordered in two lists, one of positive and the other of negative UX aspects.

3.3   Adopting the New Release of the Method

What follows describes the adoption of the new release of the method in the field, phase by phase. This comes to evaluate a state-of-the-art CAD software package developed by a well-known software house working in the engineering research and application fields. This choice comes from the availability of the product for the tests as well as that of the users to involve; moreover, it also comes from the good evaluators' knowledge about the product.

In the first phase, the 3D modeling section of the CAD package is selected as the product feature to evaluate. After that, the users are selected. For the AB and RB modules, the users must not know the selected CAD package; for the RE module, the users must know this product and have used it almost for one month. All of them are students of mechanical engineering courses with good knowledge about 3D modeling since they used one or more CAD packages in the past. In all, thirty users are involved, ten for each module. Finally, all the modules are considered in this first adoption of the new release of the method and three researchers very knowledgeable with the CAD package are selected as reviewers.

In the second phase, the guide documents for the three modules are customized respect to the specific evaluation to perform. The need is set to "generate the 3D model of the socket shown in figure. This model will be used for production. Use the CAD package available on the PC. Please respect the assigned dimensions". Figure 1 shows the drawing of the socket as reported in the guide documents. The documents for the evaluators are used as they are, while a university lab consisting of two separate rooms is selected as the environment to perform the tests. In the first room, the users generate the irMMs; in the second room, a PC with the CAD package running on it allows each user to perform the modeling activities.
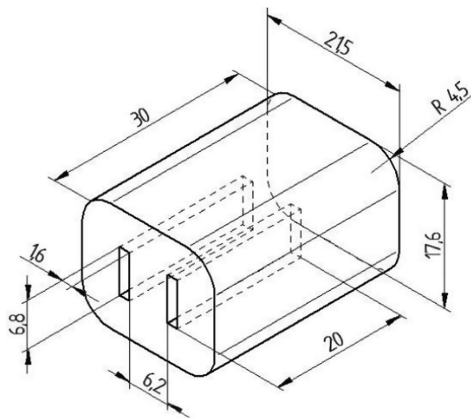

Fig. 1. The drawing of the socket to model (units are millimeters)

Once the material is ready, the tests can start (third phase). Each user associated to the RB module (hereafter, RB user) interacts freely with the CAD package for five

minutes. After that, all the users generate their irMMs. It is worth to say that, contrary to what described in the background section, during the definition of the desired meanings and emotions, the users refer them to the need and not to the specific goals. This happens because the need is so simple and immediate that it makes the same meanings and emotions come in mind for every goal. What follows are some examples of goals, meanings and emotions. One of the RE users expresses the two goals "extrusion of the main rectangle shape generated" and "cutting of the two rectangular holes done". In the RB module, a user defines four desired meanings: hobby, job, courses, manufacturing. One of the AB users defines six desired emotions: happy, excited, not engaged, tired, not relaxed and worried. Table 1 collects the excerpt of an irMM generated by one of the RB users referred to the goal "Boolean subtraction done" (please consider the first two columns only, containing the user actions and the product reactions/feedback, respectively).

Table 1. Excerpt of a user's irMM related to the goal "Boolean subtraction done" (first two columns) and evaluation of it against the real interaction (last two columns)

User actions	Product reactions/ feedback	Check	Reason and judgment
Open the menu to perform Boolean operations	The menu to select the Boolean operations is shown	Passed	
Select the command to subtract volumes		Failed	The menu command extrude is unclear. The evaluators show me that it also allows performing subtractions but its label seems to mean the opposite. Negative gap
Select the main prism	The first prism is highlighted	Passed	
Select the prism to subtract	The second prism is highlighted	Passed	
	The preview of the resulting volume is shown	Failed	I was expecting to see only the preview of the resulting volume. Nevertheless, in addition to that, the package shows also the starting prisms as shaded. This allows me verifying if the performed Boolean operation (cut) succeeded. Positive gap
Confirm the operation	The hole in the model is shown	Passed	

Once the users have generated their irMMs, they are called to evaluate the real interaction with the CAD package respect to these irMMs. By carrying on with some examples, consider again Table 1; this time, please focus on the last two columns. During the real interaction, the user highlights two gaps (the two rows with "failed" in the third column). The first one, related to the user action "select the command to

subtract volumes", refers to the command extrude and is negative; the second one, related to the product reactions/feedback "the preview of the resulting volume is shown", refers to the richer visualization of the preview of the results and it is positive. The fourth column reports the reasons for the gaps from the user's point of view and the related positive/negative judgment. Once the real interactions come to the end, the users also highlight the changes between the desired meanings and emotions and the real ones. For example, in the AB module, a user defined the desired meaning "software architecture" with the reason "easy understanding and exploiting of the mechanisms that govern the software" and with a positive judgment. After the real interaction, the judgment on the meaning becomes negative and the reason for this is "difficult understanding of the complex mechanisms in using the software". Again, two desired emotions identified by a RB user are happy and quite relaxed. After the interaction, the user changes them into strongly unhappy and not relaxed because she found too many gaps during her test.

The fourth phase generates the lists of positive and negative UX aspects for each module. The AB module generates 4 positive UX aspects and 35 negative ones; the RB module generates 8 positive and 23 negative; finally, the RE module generates 11 positive and 22 negative UX aspects. An example of positive UX aspect, coming from the RE module, is "many different file formats are available to export the model", while one negative UX aspect, coming from the AB module, is "once finished the feature extrude, the window view does not come back to isometric automatically". After that, the UX aspects generated by all the modules are collected and compared to each other to delete repetitions. In the end, the final lists of positive and negative UX aspects contain 16 and 68 items, respectively. The most popular positive UX aspect is "some menu icons are clearer and more intuitive than the classic ones (e.g., coil, hole and rib)", while the most popular negative one is "when users need to generate square holes, they must use the menu command labelled as extrude but extrusion means exactly the opposite".

3.4   First Validation of the New Release of the Method

In order to start validating the new release of the method, the UX of the same CAD package is evaluated using the old release of the method. This comes as described in the background section. Thirty users are selected; they are students coming from the same mechanical engineering courses of the users involved in the previous tests. Considering the same need, the executions of the tests and the analysis of the results highlight the two lists of UX aspects as composed by 6 positive and 52 negative.

The results of the two releases of the method can be compared thanks to the analogies between the two evaluations (same product, same number of users involved, same types of data collected, etc.). This comparison highlights that the new release of the method generated more UX aspects than the old one. This is due to the higher numbers of desired meanings and emotions, gaps and of changes of meanings and emotions in the new release. From the qualitative point of view, the most of the positive and negative UX aspects identified by the new release have been highlighted also by the old one, as for example the negative UX aspects about the feature to create square holes. This correspondence is supported also by the similarities in the desired meanings

and emotions from which several UX aspects were generated. Nevertheless, the new release of the method reports UX aspects not highlighted by the old one. These UX aspects refer to gaps (mainly negative ones) identified in specific user actions or product reactions/feedback that only users who know that product quite well can highlight. For example, the "line" command does not remain active if a specific length for the edge is keyed in and confirmed with the enter key instead of defined thanks to mouse movements and clicks only. Moreover, these UX aspects come also from the changes of meanings and emotions because the most of them are generated by RE users. This happened because the users' knowledge about the product allowed them to consider the potentialities of the product in different contexts than those where they currently use the product or used it in the past.

3.5 Identification of the Relationships Between the Results and the Evaluation Activities

The last part of the research aims at identifying relationships between the results of the evaluations and the activities that discover them. This will allow selecting the most suitable modules to adopt time by time depending on the characteristics of the results, on the resources available, on the evaluators' personal style, etc. Different cases are considered; each of them involves a different company that expresses specific requirements about the results. These requirements are translated into the language used by the method and the relationships are defined exploiting the data elaboration happened during the analysis phase. In the following, two examples of relationships are proposed.

The first relationship regards a company interested in the negative emotional impact of one of its products in order to lower it. This requirement can be translated into the consideration of those UX aspects suggested by negative changes of emotions. Analyzing the data collected for each module during the tests, the highest value of these negative changes was generated by the AB module (24), followed by the RB (6) and by the RE (2) ones. Clearly, this was determined by the presence/absence of prior knowledge about the product. AB users must not know the product; therefore, since they based their evaluation exclusively on their experiences with similar products and on what they would like to experience, the product under evaluation could likely make them disappointed. On the contrary, the users associated to the other two modules knew partially (RB) or completely (RE) what to expect from the product; for this reason, the negative emotional impacts were much lower or almost absent, respectively. Therefore, the first relationship can be formulated as "in order to evaluate the negative emotional impact of a product, the AB module is the best choice to adopt, followed by the RB and at the RE ones, in this order".

The second relationship regards a company willing to know the strong pluses about functions and procedures of one of their products in order to emphasize them further. This requirement can be translated into the consideration of the UX aspects deriving from positive gaps. Analyzing again the data collected for each module, this relationship can be formulated as "in order to evaluate the pluses of a product, the RE module is the best choice to adopt, followed by the RB and the AB modules, in this order".

It is worth to say that the relationships have been inferred from just one adoption of the new release of the method and for a specific product; therefore, they need further confirmations before taking them as granted for different products and situations.

4   Conclusions

Aiming at improving the existing irMMs-based UX evaluation method, the research described in this paper kept into consideration the users' knowledge about the product. Users already familiar with the product have been placed side by side to those who never got in touch with it before. This was achieved by patching the existing method in order to split the evaluation activities into two modules, depending on the possibility of the users to interact freely with the product before to start the tests. This made the method modular and allowed its refinement. The refining activities introduced a third module; it involves the users familiar with the product under evaluation. In order to check the improvements and validate the research results, the old and the new releases of the method were adopted to evaluate the UX of a CAD package. The comparison of the results showed that the new release generated more UX aspects, both positive and negative, demonstrating its higher completeness. In addition, the results coming from the three modules were analyzed separately and this analysis allowed highlighting relationships between the results and the modules that generated them. These relationships became suggestions about the best modules to adopt against required characteristics of the results, resources available, evaluators' personal styles, etc.

Some research perspectives could be foreseen too. First, the architecture of the modules should be made as generic as possible to allow adding/modifying them easily and quickly. Second, several activities during the tests should be optimized to avoid time-consuming tests; consequently, users would be more concentrated and the results coming from the evaluation of the real meanings and emotions would be less affected by tiredness and more reliable. Third, new adoptions of the method in the field are required to confirm the effectiveness of the improvements. These adoptions should involve different products and situations, and the needs should be more complex in order to avoid simplifications like the assignment of meanings and emotions to the whole experience - as it happened in the evaluation of the CAD package - rather than to each goal as the definition of the method would require. These further adoptions would be used also to validate the relationships identified in this research and to increase their number. Fourth, these relationships should be generalized to be available in different contexts for evaluating different products. Finally, these relationships should start to be considered not only for evaluation but also for design. In fact, several results of the method adoption, like the lists of user actions and product reactions/feedback (in other words, the irMMs as they are), could be exploited to suggest guidelines to design/redesign products rather than simply to evaluate them. All of this could be the starting point for making the irMMs-based method turn to a UX design aid.

References

1. Law, E.L.C., Roto, V., Hassenzahl, M., Vermeeren, A.P.O.S., Kort, J.: Understanding, scoping and defining user experience: a survey approach. In: Conference on Human Factors in Computing Systems, CHI 2009, pp. 719–728 (2009). https://doi.org/10.1145/1518701.1518813

2. Park, J., Han, S.H., Kim, H.K., Cho, Y., Park, W.: Developing elements of user experience for mobile phones and services: survey, interview, and observation approaches. Hum. Factor. Ergon. Manuf. Serv. Ind. **23**(4), 279–293 (2013). https://doi.org/10.1002/hfm.20316

3. Brave, S., Nass, C.: Emotion in human-computer interaction. In: Sears, A., Jacko, J.A.: The Human-Computer Interaction Handbook, pp. 81–96. CRC Press (2012)

4. Gentner, D., Stevens, A.: Mental Models. Lawrence Erlbaum Associates, Hillsdale (1983)

5. Greca, I.M., Moreira, M.A.: Mental models, conceptual models and modelling. Int. J. Sci. Educ. **22**(1), 1–11 (2000). https://doi.org/10.1080/095006900289976

6. Rauschenberger, M., Pérez Cota, M., Thomaschewski, J.: Efficient measurement of the user experience of interactive products. How to use the user experience questionnaire (UEQ). Example: spanish language version. Int. J. Artif. Intell. Interact. Multimed. **2**(1), 39–45 (2013). https://doi.org/10.9781/ijimai.2013.215

7. Bernhaupt, R., Pirker, M.: Methodological challenges of UX evaluation in the living room: developing the IPTV-UX questionnaire. In: 5th Workshop on Software and Usability Engineering Cross-Pollination, PUX2011, pp. 51–61 (2011)

8. Burmester, M., Mast., M., Kilian, J., Homans, H.: Valence method for formative evaluation of user experience. In: Designing Interactive Systems Conference 2010, DIS 2010, pp. 364–367 (2010). https://doi.org/10.1145/1858171.1858239

9. Filippi, S., Barattin, D.: User experience (UX) evaluation based on interaction-related mental models. In: Ahram, T., Falcão, C. (eds.) AHFE 2017. AISC, vol. 607, pp. 634–645. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-60492-3_60

10. Kujala, S., Vogel, M., Pohlmeyer, A.E., Obrist, M.: Lost in time: the meaning of temporal aspects in user experience. In: 13th Computer-Human Interaction Conference CHI, Extended Abstracts on Human Factors in Computing Systems, pp. 559–564 (2013). https://doi.org/10.1145/2468356.2468455

11. Karapanos, E., Zimmerman, J., Forlizzi, J., Martens, J.B.: Measuring the dynamics of remembered experience over time. Interact. Comput. **22**, 328–335 (2010). https://doi.org/10.1016/j.intcom.2010.04.003

12. Sng, K.H.E., Raviselvam, S., Anderson, D., Blessing, L., Camburn, B.A., Wood, C.: A design case study: transferring design processes and prototyping principles into industry for rapid response and user impact. In: 21st International Conference on Engineering Design, ICED17, pp. 349–358 (2017)

13. Norman, D.: The Design of Everyday Things: Revised and Expanded Edition. Basic Book, New York (2013)

14. Ullman, D.G.: The Mechanical Design Process, 4th edn. McGraw-Hill, New York (2010)