A Method and Tool for Analyzing Usability Testing Data in the Business Sector

Panagiotis Germanakos[1(✉)] and Ludwig Fichte[2]

[1] SAP SE, Dietmar-Hopp-Allee 16, 69190 Walldorf, Germany
`panagiotis.germanakos@sap.com`
[2] Shopify Inc., 150 Elgin Street, Ottawa, ON K2P 1L4, Canada
`ludwig.fichte@shopify.com`

Abstract. Usability testing is a critical phase of the User-Centered Design and development process, where product teams can observe and measure the usability of their software solutions' functionality, interactions and user interfaces. The subsequent validation sessions may generate a large, overwhelming, volume of empirical and unstructured data that need to be analyzed to determine meaningful outcomes for informed decisions. In this paper, we overview a new method and tool, namely EUREKA, guiding a product team, through a real-life business scenario, to analyze its usability study qualitative data and to produce comparable and replicable results, improving the User Experience of its product.

Keywords: User Experience · User-Centred Design ·
Usability Testing · Data analysis · Framework

1 Introduction

The current business landscape could be characterized by a rapid digital transformation of products and services, where User Experience (UX) is in the center of attention. Business organizations invest a significant amount of resources on related activities for enhancing the quality and "look-and-feel" of their solutions [9], for gaining the competitive advantage in the market. Overarching goal is to offer more usable interactions with their business processes through the user interfaces that will increase the desirability, accessibility, usefulness, etc., of their functional systems. Such qualities can be evaluated through a series of usability tests during the validation phase of the User-Centred Design and software development process [7]. Main concern is to thoroughly analyze the collected information, behaviours, and observations and facilitate an inclusive understanding of the vague and unstructured data transformation into actionable items. Accordingly, the project teams will be able to make informed decisions, increasing the, e.g., effectiveness, efficiency and satisfaction [6,8] of their end-users during task

execution. In this respect, we overview a new method and tool that guides the analysis of collected usability testing qualitative data, namely EUREKA (Engineering Usability Research Empirical Knowledge & Artifacts), providing a balanced and semantically enriched qualitative and quantitative perspective of the outcome to the benefit of the product teams and the end-users.

2   An Overview of the EUREKA Workflow Through a Real-Life Business Scenario

EUREKA is an end-to-end Workflow-as-a-Service methodology (see Fig. 1) and tool (currently in the form of an .xls prototype) for analyzing usability testing feedback (see also [2,3] for more details). It may be regarded as an improved approach that could provide guidance through a highly synergetic environment during the analysis of the empirical data captured from the validation sessions. It consists of 4 main goal-directed phases, i.e., Discover, Learn, Act, and Monitor, that may embody concepts (e.g., as interrelated layers) of a taxonomy that represents a comprehensive paradigm for supporting (either as a guide or as standalone categories and classifications) the extraction of insightful learning outcomes and meaningful action items through one or more refinement cycles during the qualitative data analysis process.
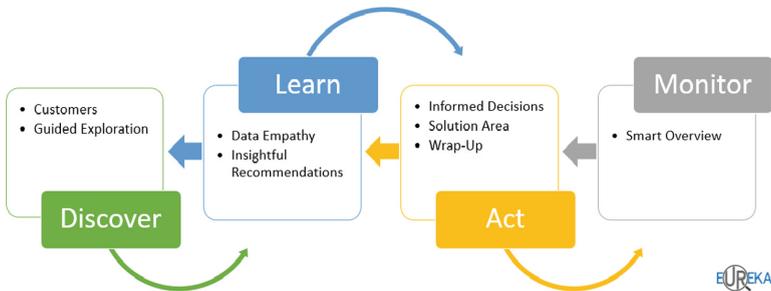
Fig. 1. Main process steps of EUREKA workflow

Next we briefly describe the main process steps of the EUREKA workflow through a real-life business scenario. A product team, composed of professionals with different roles (e.g., Product Owner, Business Expert, User Researcher, Interaction Designer, Architect, Developer), sharing different knowledge, experiences, expectations and data analysis skills, returned to their headquarters after executing a series of on-site usability testing sessions with their customers. Main scope was to validate their functional prototype of a newly launched mobile application. They tested the product with 3 end-users from 4 different customers that they visited in total. They now want to analyze and gain an inclusive insight on the empirical data and observations they collected (in total 320 feedback items) so to make viable decisions for their priorities and the backlog development

items. However, they found themselves to be overwhelmed from the excessive amounts of unstructured usability test data (e.g., opinions, suggestions, sentiments, experiences, etc.) and for not knowing how to make a start; so to assign structure and meaning within a reasonable amount of time. Hence, they decided to use EUREKA, for guiding them through the data analysis process.

The team begins the analysis with the **Discover** phase, formulating a first understanding of the collected data, by capturing, consolidating, synthesizing and iterating on the content of the raw seeds of information. This stage consists of two subsequent process steps: The *Customers*, where the team records descriptive details about the end-users as well as the degree that they fit to the expected user profile and tasks. The role fit will be used throughout the analysis for weighting the responses of the end-users, assigning importance and treating data with fairness. During the *Guided Exploration* process step the team applies various iterative operations on the raw data for identifying inconsistencies, gaps or misconceptions, turning them into a coherent data set of feedback items. This activity will produce optimized clusters with the end-users' weighted references assigned to each one of them, specific aggregation calculations and the success with assistance for each end-user (degree of external influence for accomplishing a task). Moving to the next phase, **Learn**, the team can assign meaning to the derived clusters by identifying their type, polarity, association with other artefacts (e.g., use cases, designs), usability issues [4] or recommendations that they express, and their relevant importance and impact on the given application. Those actions are taking place during the *Data Empathy* subsequent process step that facilitates the objective observation and unbiased interpretation of the data by the team, as well as the precise association of the various units of information in an attempt to reveal holistic and inclusive meanings of a feedback item (or cluster). In parallel, the *Insightful Recommendations* process step generates suggestions (by considering the relevant importance and impact) if the team should take an immediate action or not. The following phase, **Act**, encompasses 3 subsequent process steps: The *Informed Decision*, that the team can discuss and assign possible high-level alternative solutions on one or more usability issues, and take actions as regards, i.e., proceed with one solution or not, and tracking their progress. The team may also deep dive to their solutions through the *Solution Area* process step, exploring alternative approaches and their viability for solving the discovered usability issues. They can determine the solutions' effectiveness and coverage across the usability issues, estimated effort, calculated risk, and likelihood of timely completion (based on the assigned effort and risk). Lastly, through the *Wrap-Up* process step, the team can similarly analyze all the information collected from post-questions (e.g., impressions, improvement points, or situation-specific comments), as well as usability or UX test tools' responses for cross-evaluation of the main tasks (e.g., SUS [1], or UEQ [5]). Finally, the product team can benefit from the **Monitor** phase, referring to the *Smart Overview* subsequent process step, that facilitates the continuous monitoring and exploration of the information that has been analyzed in the previous phases. The team can create visually enhanced cards containing various statis-

tics for the feedback items' clusters, issues judgment, tasks assistance, usability issue types, etc., that may provide a quick overview of the empirical research outcome. It can also provide a structured documentation, fast reporting, and guide for prioritizing actions and decisions through an informed drill-down on the reformulated semantic data (e.g., by applying filters).

3   Preliminary Evaluation Results

During the year 2018 we evaluated EUREKA using 7 different data sets (approx. 1300 feedback items) from usability tests and 45 end-users, with different business roles and data analysis expertise. During the data analysis sessions, we observed the application of the method and tool, we conducted interviews and formulated focus groups gathering constructive feedback for its use and value.
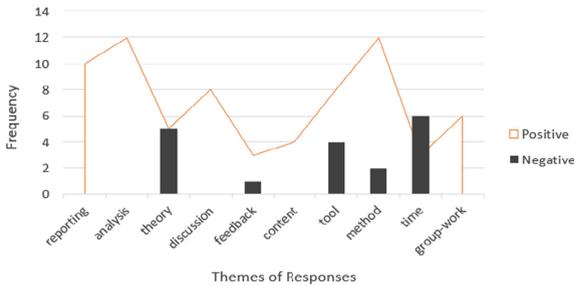
Fig. 2. Frequency of themes across the participants' responses (Color figure online)

For the analysis we applied: (a) frequency of themes for synthesizing the collected open-ended responses (referring to impressions, challenges, suggestions, etc.), and (b) sentiment analysis classifying them into two categories regarding their polarity, i.e., positive or negative. In Fig. 2, we highlight the main themes considering how often they semantically reappeared in each feedback statement across the participants. Indicatively, we observe that for the positive (orange line) the "analysis", "method", and "tool" appeared proportionally more frequently to the responses of the participants, while for the negative (grey bar) the "time" and "theory". For the former, example representative statements include: "None of the feedback items got lost during the analysis, and they are all relevant to the final assessment", "the method itself is of great help to UX professionals and development teams", or "I was amazed how I received numbers instead of only text as an outcome using the tool"; while for the latter: "We would need some time to familiarize ourselves with the prototype tool", or "the theory behind turns out to be a bit complex in cases".

4   Conclusion

EUREKA method and tool encapsulates a comprehensive and rather flexible approach that facilitates the analysis of qualitative data collected from usability tests. Its added value focuses on the collaborative transformation of vague feedback contents and behaviours into insightful and actionable items which can be tracked, followed-up and lead to informed decisions for business professionals that they want to enhance the UX and quality of their products.

References

1. Brooke, J., et al.: SUS-A quick and dirty usability scale. Usability Eval. Ind. **189**(194), 4–7 (1996)
2. Geramanakos, P., Fichte, L.: #EUREKA - Create Data Empathy Not Melancholy (2018). https://experience.sap.com/skillup/eureka-create-data-empathy-not-melancholy/. Accessed on 07 Apr 2019
3. Germanakos, P., Fichte, L.: EUREKA: engineering usability research empirical knowledge and artifacts. In: Zaphiris, P., Ioannou, A. (eds.) LCT 2018. LNCS, vol. 10924, pp. 85–103. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91743-6_6
4. Keenan, S.L., Hartson, H.R., Kafura, D.G., Schulman, R.S.: The usability problem taxonomy: a framework for classification and analysis. Empirical Softw. Eng. **4**, 71–104 (1999)
5. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (ed.) USAB 2008. LNCS, vol. 5298, pp. 63–76. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-89350-9_6
6. Nielsen, J.: Usability Engineering. Elsevier, Amsterdam (1994)
7. Preece, J., Rogers, Y., Sharp, H.: Interaction Design: Beyond Human-Computer Interaction. Wiley, Hoboken (2015)
8. Shneiderman, B.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. Pearson Education India (2010)
9. UserTesting: 2017 UX and User Research Industry Survey Report (2017). https://info.usertesting.com/ux-industry-survey-2017.html. Accessed 20 Nov 2018