Identifying Phenotypes & Genotypes: A Case Study Evaluating an In-car Navigation System

Georgios Papatzanis

Introduction

Various techniques can be used for the evaluation of interactive systems. Techniques are classified according to their approaches in conducting evaluation: *analytically* when a simulation is performed by an expert/analyst to predict the behaviour of the user and detect potential problems, without the involvement of users; *empirically* when the system is tested by users while their performance and problems are recorded.

The aim of this paper is to report on a study where both analytical and empirical approaches were employed to evaluate an in-car navigation device. In this study we concentrated solely on tasks related to the programming of the device (destination entry) before the user starts driving the car. We look at the results from a qualitative perspective; we do not seek to establish efficiency counts (number of usability problems) for different techniques or approaches. Instead we analyse the results produced by the two classes of approach – analytical and empirical – and compare them in terms of their diversity and the insight they provide to the analyst in respect of the overall usability of the system and its potential improvement. We investigate the variance of results between the classes of approach and explore the association of genotypes and phenotypes with the empirical and analytical classes of approach respectively.

Method

The case study was executed in two discrete parts: analytical and empirical evaluation, followed by an analysis of the results comparing the two classes of approach. In each, the usability of the selected application was assessed against a predefined scenario and set of tasks. The scenario and tasks were based on the activities carried out by the driver prior to driving to the destination, i.e., preparing an itinerary on the navigation device. Such tasks take place in the car while stationary. This set of tasks enabled us to assess a wide range of primary functions that are frequently used in such devices.

In the analytical part of the study we applied a series of analytical methods in the evaluation of the navigation system. Four methods were chosen for the analytical part of the study, employing a diverse approach to evaluation. The methods selected are characterised by a varying degree of formality, with each advocating a different approach to user interface evaluation. Each method has its own potential merits in the evaluation of this device. Cognitive Walkthrough [6] was selected as it is most suitable for walk-up-and-use interfaces. EMU (Evaluating Multi-Modal Usability) [4] is a technique specifically implemented for multimodal systems, thus appropriate for this type of device. UAN (User Action Notation) [1] provides an extensive notation, incorporating temporal issues and patterns for the specification of the interface. Leveson's design guidelines [5] were selected because of their focus on error detection and analysis. The diversity of these techniques gives us an increased capacity for the detection of usability issues, giving a wide range to compare against those found empirically.

In the second part of the study we carried out an empirical evaluation of the device, using the same scenario and tasks as in the first part of the study. The empirical evaluation was carried out in a usability laboratory, as the context (being in a car) is not relevant for the set of tasks selected for this study. We focused our attention on the usability issues that drivers encounter in the use of such devices, employing an exploratory approach.

Results

The types of issues captured by analytical and empirical techniques vary significantly. Some usability problems were identified by both classes of approach (analytical and empirical), but many were identified only by one or the other. In the previous section we presented a set of usability problems representing these categories and as tokens of the usability problems identified.

An important aspect that emerges when looking at the results is the variability between the coverage of results reported by analytical and empirical approaches. There is only a small overlap on the issues

identified by the two approaches. The vast majority of usability problems were independently identified by one class of approach only.

Under closer investigation we also observe that the type of problems detected by the approaches is significant. While the analytical techniques identified mainly usability problems that might create difficulties to the users, the empirical data demonstrated specific instances of user behaviour where users experienced such difficulties. The usability problems reported by the empirical approach are associated with the manifestations of user errors, while the usability problems reported by the analytical approach correspond to the underlying cause of such manifestations. This correspondence thus relates to the phenotype – observable manifestations of an incorrectly performed action – and the contrasting genotype – the underlying likely cause [2] [3].

In order to investigate the association of genotypes and phenotypes with their respective classes of approach – empirical and analytical, the first author and a further two usability experts independently assessed the issues identified in the study in terms of genotypes and phenotypes. The experts were able to match the majority (over 95%) of the issues to the type of error, as we had hypothesised with the correlation between genotypes, phenotypes and their respective classes of approach. The issues identified by the empirical class of approach were assigned as genotypes, whereas the issues identified by the analytical class of approach were assigned as phenotypes.

Conclusion

In this study we set out to compare different evaluation techniques by evaluating the usability of a car navigation device. Our efforts were concentrated on the aspects of the device relating to the preparation of a route, before the device commences with the navigational instructions to the driver of the car.

We examined the analytical and empirical techniques that were employed during the study. Each technique employed in this study offers a different perspective into the usability evaluation of interactive systems and identified different sets of issues. In this study we focused on the kind of usability problems reported from each class of approach. According to the results of the study, the analytical class of approach is most powerful as a way of identifying genotypes, while the empirical class of approach is best at identifying phenotypes. These results support the argument that a combination of analytical and empirical approaches can offer a richer insight into the usability of the system and give the usability practitioner greater argumentative power, as their findings complement each other.

The association of phenotypes with their respective genotypes is a difficult task, but necessary in the process of increasing the usability of a system, when adopting such an approach. Taxonomies identifying domain specific genotypes and phenotypes could eventually assist the analyst relating observational behaviour to underlying cause, resulting in a deeper insight into the usability of a system.

In order to assess further the scope of each technique and approach in a dynamic environment, we are carrying out another study where the tasks selected are representative of the user experience while driving and taking instructions from a navigation device. This future study will give us further insight into the appropriateness of the methods when using such devices in a constantly changing environment and where the goals of the users are not preconceived as is the case in this study.

Acknowledgments. The work described in this paper has been funded by the EPSRC Human Error Modelling project (GR/S67494/01 and GR/S67500/01). We are grateful to all participants in our studies.

References

1. Hartson, H.R. and Gray, P.D., *Temporal Aspects of Tasks in the User Action Notation.* Human-Computer Interaction, 1992. **7**(1): p. 1-45.
2. Hollnagel, E., *The phenotype of erroneous actions.* International Journal of Man-Machine Studies, 1993. **39**(1): p. 1-32.
3. Hollnagel, E., *Cognitive Reliability and Error Analysis Method*. 1998, Oxford: Elsevier Science Ltd.
4. Hyde, J.K., *Multi-Modal Usability Evaluation*. 2002, PhD thesis. Middlesex University.
5. Leveson, N.G., *Safeware: System Safety and Computers*. 1995: Addison-Wesley.
6. Wharton, C., Rieman, J., Lewis, C., and Polson, P., *The cognitive walkthrough method: a practitioner's guide*, in *Usability inspection methods*. 1994, John Wiley & Sons, Inc. p. 105-140.