COMP2213 - Interaction Design

Reflection 4: MySouthampton Mobile Applicaiton Heuristic Evaluation Workshop

Daniel Braghis
ddb1u20@soton.ac.uk
UoS ID: 32161204
14/12/2022

This week's workshop session addressed the final step of the design process – evaluation. Our design team was tasked with picking a mobile application which we frequently use and switch roles from application users to HCI experts applying usability heuristic evaluation to identify, classify and propose solutions to design flows. MySouthampton was the application we have chosen to perform summative evaluation on (the application development is completed).

As we already know, heuristic evaluation is divided in three stages: briefing session, evaluation period and debriefing session (Sharp et al., 2019, p. 441) [1]. In our task description these were equivalent to preparing a task list appropriate for the application, following the task list and identifying heuristic violations, and finishing with a severity assignment discussion of the issues found. As for the usability heuristics used, the course materials suggest Shneiderman's and Nielsen's (Sharp et al., 2019, p. 437,445) as the main sets of principles but there are various interpretations that can be adapted to the specific needs of the design team, a practice encouraged in the field (Sharp et al., 2019, p. 445). Such flexibility to me is an indicator of the practicality the methodologies are trying to achieve, forgoing a prescriptive structure in favour of covering the needs of the designers from various industries. In our case, as a starting point, Nielsen's heuristics guided the evaluation.

Our analysis revealed several issues in the MySouthampton app: a critical error that failed to authenticate the user accompanied by a cryptic error message and other inconsistencies in app design/labeling that might confuse the users or offer a subpar UX experience. However, we have also identified a set of violations of heuristics that where questionable like the application embedding or linking to already existent systems. According to Nielsen heuristics these are at least partly in violation of the "Recognition Rather than Recall" heuristic (Sharp et al., 2019, p. 437) but our group questioned if that would be an issue for a service integrator like MySouthampton forgoing the consistency to better serve the users system interfaces that might be already familiar to them and avoid "reinventing the wheel" by integrating them completely into the application. This led us to doubt the efficiency of the method and wonder how much we have missed by concentrating on fitting the design issues found to a specific heuristic.

Both the course book and research ascertain that "the experts tend to identify more false alarms and miss more problems than they have true hits" (Sharp et al., 2019, p. 443) and "does little to guide analysis towards hard to find problems" failing to predict 43% of severe problems of varying frequency and introduces a significant amount of false positives that impact negatively the credibility of the method and expert (Cuckton et al., 2001). This evidence builds a strong case against heuristic evaluation as a replacement to user studies but for educational purposes it is a time inexpensive solution to identify surface level problems. Our design team agreed that in a professional setting we would avoid heuristic analysis mainly due to concerns around costly false positives. We are giving preference to more time and resource consuming A/B testing, questionnaires, or laboratory user observation we have learned about in the lectures that I assume result in a more realistic evaluation of the user experience.

Following the lecture on Values and Interaction, I have noticed that none of the Nielsen's heuristics addressed the user values during evaluation which I find alarming for a product like MySouthapton that has to satisfy the needs of a diverse student body including the LGBTQ+ community or individuals with accessibility needs. As a designer responsible for the app, I would add a heuristic checking for accessibly options available in the interface and one addressing symbols and wordings that are user background agnostic (e.g., using simple language or icons that are universally accepted standards). Suggested research on the topic also recommends "Value-oriented Mockup, Prototype or Field Deployment" (Batya et al., sec. 2.9) as a viable way to elicit value centered user feedback on a low fidelity mockup.

Finally, the workshop, given the time and resource constraints, succeeded at drawing our attention to the advantages and tradeoffs of the heuristic design evaluation. However, as a practical tool, the narrow scope it addresses (neglecting accessibility and values) coupled with the serious concerns around false positives (Cuckton et al., 2001) discussed earlier, lead me to conclude that this evaluation methodology requires high familiarity with the product and expert understanding of the HCI field to be successfully employed in practice. As novice interaction designers like ourselves, given the appropriate resources (especially time), preference should be given to methods that don't rely so heavily on our (limited) experience.

---

1. Interaction Design: Beyond Human-Computer Interaction, Helen Sharp, Fifth edition, 2019,Wiley Indianapolis, IN, 1119547350

2. Cockton, G., Woolrych, A., Understanding Inspection Methods: Lessons from an Assessment of Heuristic Evaluation. In: Blandford, A., Vanderdonckt, J., Gray, P. (eds) People and Computers XV-Interaction without Frontiers, 2001, Springer, London, https://doi.org/10.1007/978-1-4471-0353-0_11

3. Foundations and Trends in Human-Computer Interaction, Batya Friedman, David G. Hendry, Alan Borning, Volume 11, Issue 2, 22 11 2017, pp 63–125, https://doi.org/10.1561/1100000015