

Mobile Application Usability: Heuristic Evaluation and Evaluation of Heuristics

Ger Joyce, Mariana Lilley, Trevor Barker and Amanda Jefferies

Abstract Many traditional usability evaluation methods do not consider mobile-specific issues. This can result in mobile applications that abound in usability issues. We empirically evaluate three sets of usability heuristics for use with mobile applications, including a set defined by the authors. While the set of heuristics defined by the authors surface more usability issues in a mobile application than other sets of heuristics, improvements to the set can be made.

Keywords Human factors · Usability · Mobile apps · Heuristic evaluation

1 Introduction

Human-Computer Interaction (HCI) researchers and practitioners use traditional usability evaluation methods to evaluate the usability of mobile applications. Yet, these traditional methods do not always consider applications built for small screens and rapidly changing environments [1, 2]. As mobile application use has grown exponentially in recent years [3], HCI researchers and practitioners need to address this issue [4]. As argued by the authors in previous work [5, 6], popular usability evaluation methods, such as a Heuristic Evaluation [7], may be modified for use with mobile applications. In this paper, the authors empirically investigate that

G. Joyce (✉) · M. Lilley · T. Barker · A. Jefferies
School of Computer Science, University of Hertfordshire,
College Lane, Hatfield, Hertfordshire AL10 9AB, UK
e-mail: gerjoyce@outlook.com

M. Lilley
e-mail: m.lilley@herts.ac.uk

T. Barker
e-mail: t.i.barker@herts.ac.uk

A. Jefferies
e-mail: a.l.jefferies@herts.ac.uk

claim. This work will be of importance to HCI practitioners, educators, and researchers—indeed any teams that focus on developing and evaluating the usability of mobile applications.

2 Related Work

Expert-based usability inspection methods, whereby a group of HCI experts evaluates a user interface against a set of principles are currently well established. In particular, Heuristic Evaluation is widely known for being fast and inexpensive [8], as well as for its ability to find more usability problems when compared to other methods [9]. Despite an argument that Heuristic Evaluation may not be as effective as it claims [10], the method is used quite extensively.

As mobile devices become more popular, HCI researchers and practitioners can use Nielsen's popular set of heuristics to evaluate the usability of mobile applications. However, several researchers have argued for the modification of Nielsen's heuristics in order for a more effective usability evaluation of mobile applications [11, 12]. Consequently, since 2003 researchers have defined several sets of guidelines to evaluate the usability of mobile User Interfaces [13–16]. Unfortunately, this research has not addressed vital issues within the mobile phenomena, such as rapidly changing environments, the potential of mobile devices to reduce user's workloads, and the importance of First Time User Mobile Experience. Instead, these works have focused on other areas, such as the ergonomics of a mobile device, and how to find a mobile device if lost.

3 Approach

Our approach within this study was twofold:

1. A Heuristic Evaluation of a mobile application using three sets of heuristics;
2. An Evaluation of Heuristics following the Heuristic Evaluation using a survey.

One of the sets of heuristics was a modified version of a set previously defined by the authors [5]. This set of heuristics account for areas vital to the mobile phenomena, including rapidly changing environments, the potential of mobile devices to reduce user's workloads, and the importance of First Time User Mobile Experience. Other than the set of heuristics from the authors, we selected two other sets of heuristics for the study; namely Nielsen [17] and Bertini et al. [15]. The reason behind this selection was that Nielsen's is one of the most popular sets of heuristics today, and Bertini et al. defined their set for mobile devices.

The authors recruited six HCI Experts using purposive sampling (4 Female, 2 Male). Participants had between 1 and 20 years within HCI experience

(Mean = 7.5, SD = 6.9), and between 0 and 6 years experience within Mobile HCI (Mean = 2.91, SD = 2.2). The study was conducted between February 26th, 2015 and March 16th, 2015. While a small number of participants, the number recruited by the authors was greater than Nielsen's recommendation of three to five evaluators [18]. To reduce the possibility of bias within the study, we assigned a letter to each set of heuristics, and counterbalanced the order of heuristics. Consequently, participants did not know which set of heuristics the authors had defined. Additionally, many aspects of the study were controlled, including the mobile device, mobile application, and the environmental conditions within which the study was conducted.

3.1 Tasks

In a within-subjects study, six participants ($n = 6$) completed three tasks each on a travel app from a well-established provider. Participants attempted each task on an LG G2 running Android 4.4.2 under good lighting and low ambient noise conditions, as would be expected in a Usability Testing lab. The tasks were:

1. Find a hotel near your current location using GPS for one adult that is available within the next two weeks.
2. Find a return flight for one adult in economy class from London Heathrow to Paris.
3. Read a review of a restaurant in the UK, marking the review as helpful.

3.2 Mobile Application Heuristics

The mobile applications usability heuristics modified from previous work from the authors [5] are below. We designed these with SMART (short for Smartphone Heuristics) to differentiate the heuristics from other sets.

SMART1 Provide immediate notification of application status. Ensure the mobile application user is informed of the application status immediately and as long as is necessary. Where appropriate do this non-intrusively, such as displaying notifications within the status bar.

SMART2 Use a theme and consistent terms, as well as conventions and standards familiar to the user. Use a theme for the mobile application to ensure different screens are consistent. Also create a style guide from which words, phrases and concepts familiar to the user will be applied consistently throughout the interface, using a natural and logical order. Use platform conventions and standards that users have come to expect in a mobile application such as the same effects when gestures are used.

SMART3 Prevent problems where possible; Assist users should a problem occur. Ensure the mobile application is error-proofed as much as is possible. Should a problem occur, let the user know what the problem is in a way they will understand, and offer advice in how they might fix the issue or otherwise proceed. This includes problems with the mobile network connection, whereby the application might work offline until the network connection has been re-established.

SMART4 Display an overlay pointing out the main features when appropriate or requested. An overlay pointing out the main features and how to interact with the application allows first-time users to get up-and-running quickly, after which they can explore the mobile application at their leisure. This overlay or a form of help system should also be displayed when requested.

SMART5 Each interface should focus on one task. Being focusing on one task ensures that mobile interfaces are less cluttered and simple to the point of only having the absolute necessary elements onscreen to complete that task. This also allows the interface to be glanceable to users that are interrupted frequently.

SMART6 Design a visually pleasing interface. Mobile interfaces that are attractive are far more memorable and are therefore used more often. Users are also more forgiving of attractive interfaces.

SMART7 Intuitive interfaces make for easier user journeys. Mobile interfaces should be easy-to-learn whereby next steps are obvious. This allows users to more easily complete their tasks.

SMART8 Design a clear navigable path to task completion. Users should be able to see right away how they can interact with the application and navigate their way to task completion.

SMART9 Allow configuration options and shortcuts. Depending on the target user, the mobile application might allow configuration options and shortcuts to the most important information and frequent tasks, including the ability to configure according to contextual needs.

SMART10 Cater for diverse mobile environments. Diverse environments consist of different types of context of use such as poor lighting conditions and high ambient noise are common ailments mobile users have to face every day. While the operating system should allow the user to change the interface brightness and sound settings, developers can assist users even more for example by allowing them to display larger buttons and allowing multimodal input and output options.

SMART11 Facilitate easier input. Mobile devices are difficult to use from a content input perspective. Ensure users can input content more easily and accurately by, for instance displaying keyboard buttons that are as large as possible, as well as allowing multimodal input and by keeping form fields to a minimum.

SMART12 Use the camera, microphone and sensors when appropriate to lessen the user's workload. Consider the use of the camera, microphone and sensors to

lessen the users' workload. For instance, by using GPS so the user knows where they are and how to get there they need to go, or by using OCR and the camera to digitally capture the information the user needs to input, or by allowing use of the microphone to input content.

3.3 Severity Ratings

The usability issue severity ratings used for this study were adapted from Sauro [19]:

- *Minor*: Causes some hesitation or irritation
- *Moderate*: Causes occasional task failure for some users or causes delays and moderate irritation
- *Critical*: Leads to task failure or causes extreme irritation.

4 Results

The evaluators found 145 usability issues (Mean = 48, SD = 9) (Fig. 1). Each evaluation took approximately three hours, with the subsequent analysis taking two days.

While Bertini et al. had defined their set of heuristics for mobile devices, if not specifically for mobile applications, this set surprisingly did not find as many usability issues as Nielsen's or the SMART mobile heuristics we had defined. Nielsen's heuristics, being quite generic and designed for general user interfaces,

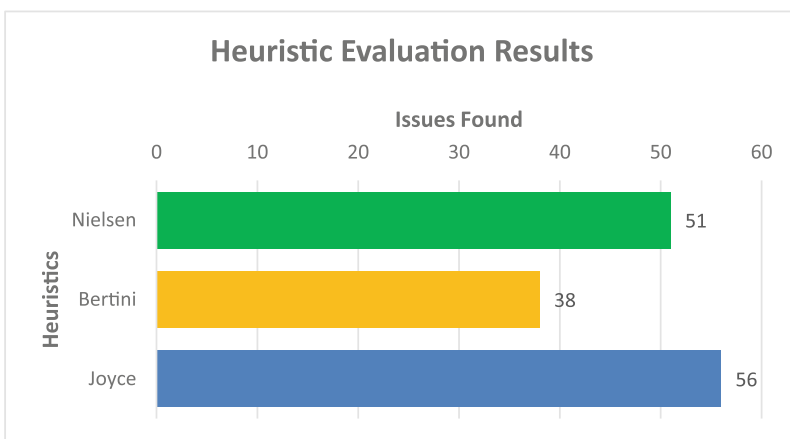


Fig. 1 Heuristic evaluation results

scored quite well. However, our SMART heuristics found the most usability issues, including critical issues.

Following the Heuristic Evaluation, each participant answered several survey questions and offered free text comments to evaluate the same sets of heuristics. This approach gave further insight into the potential for participants to use the SMART heuristics in a professional context, or if changes were required. The questions asked, and the subsequent results, follow.

- Survey Question 1. I would be confident in using this heuristic set to evaluate usability within mobile applications in a professional context.

Creating a set of heuristics applicable to any domain is part of the challenge. Ensuring that the HCI community use a set of heuristics is also part of this challenge. Therefore, we asked participants to what extent they would agree or disagree that they would be confident in using each set of heuristics to evaluate the usability of mobile applications within a professional context. Both Nielsen's and SMART heuristics scored well, with the heuristics from Bertini et al. not scoring as well (Fig. 2).

If a set of heuristics is difficult-to-use, learn, or understand, the HCI community may use other evaluation methods, potentially those that find fewer usability issues. To that end, the next set of survey questions focused on ease-of-use, learning and understanding:

- Survey Question 2: I felt the set of heuristics were easy-to-use.
- Survey Question 3: I felt the set of heuristics were easy-to-learn.
- Survey Question 4: I felt the set of heuristics were easy-to-understand.

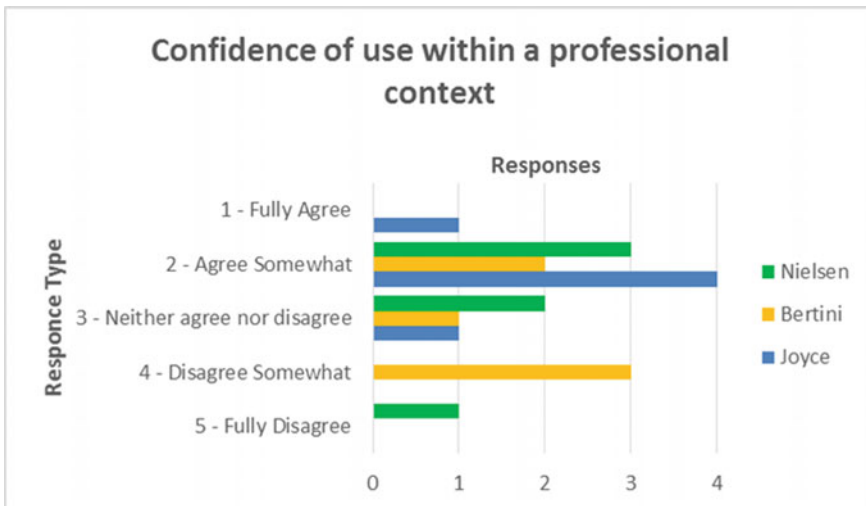


Fig. 2 Participants' confidence in using each heuristic set to evaluate the usability of mobile applications within a professional context

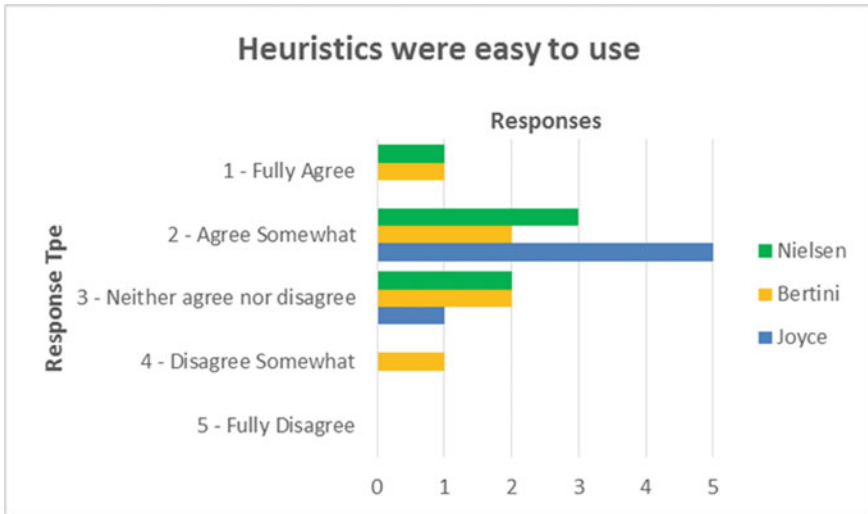


Fig. 3 Participants' perception towards ease-of-use of each set of heuristics

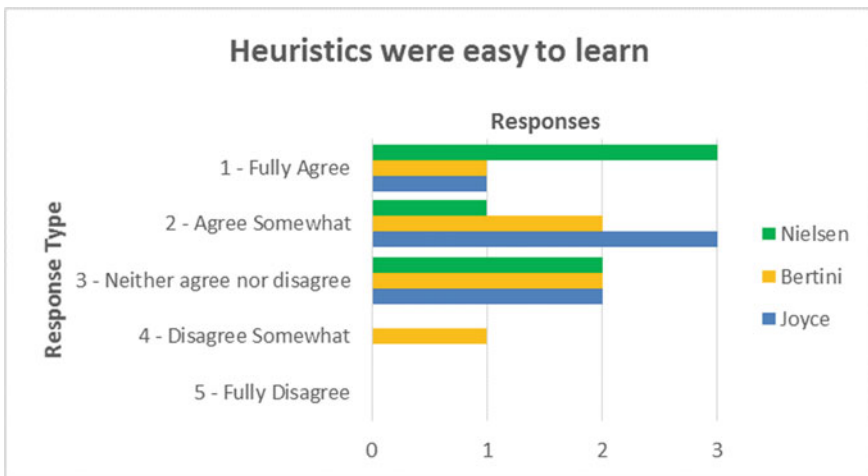


Fig. 4 Participants' perception towards ease-of-learning of each set of heuristics

Regarding ease-of-use, our heuristics scored well overall. Yet, none of the participants fully agreed that our heuristics were the easiest to use (Fig. 3). In terms of ease-of-learning, participants considered Nielsen's heuristics to be easier to learn than other set of heuristics. This is possibly due to familiarity as many HCI practitioners use Nielsen's heuristics regularly. Following Nielsen's heuristics, our heuristics scored higher than Bertini's (Fig. 4). Regarding ease-of-understanding, both Nielsen's and the SMART heuristics from the authors scored identically, with the heuristics from Bertini et al. trailing behind (Fig. 5).

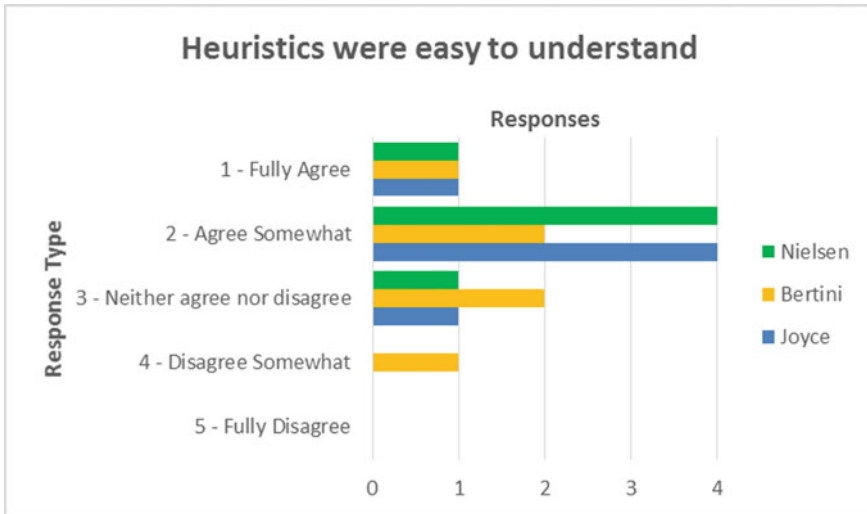


Fig. 5 Participants' perception towards ease-of-understanding of each set of heuristics

5 Analysis

The number of usability issues found during the Heuristic Evaluation differed for all three sets of heuristics. Overall, Nielsen's heuristics scored quite well, most likely because this set of heuristics is generic and applicable to most types of user interface. Conversely, the heuristics from Bertini et al. did not score as well. There could be a number of reasons for this; for instance, this set of heuristics focused on a number of areas that are not relevant to most mobile applications, such as the findability of the mobile device.

Between both the Heuristic Evaluation and Evaluation of Heuristics phases of this study, the authors set of SMART heuristics scored higher than the sets of heuristics from Nielsen and Bertini et al. in almost all areas. Not only did the SMART heuristics find the most usability issues, participants also perceived the SMART heuristics as being the most applicable for mobile application usability evaluations. Comments from participants reflected this perception:

- “P2: Set C (Joyce et al.) covers essential evaluations for mobile applications.”
- “P4: Heuristic A (Nielsen) is too broad to apply to the mobile experience. This is a strong foundation for the categories that need to be evaluated, however the guidelines need to be tweaked to cater to specific needs of mobile users.”

Interestingly, while participants found that the heuristics from Bertini et al. were applicable to mobile, participants commented that the wording on the heuristics and descriptions was “a bit clunky (P4)”.

However, while the SMART heuristics from the authors scored highly in most areas, they fell behind Nielsen's in two areas, namely ease-of-use and ease-of-learning. Reviewing participants' comments will help to understand how we can improve the SMART heuristics further:

- P1: "...decrease the number of principles and offer a similar completeness."
- P1: "The description for each heuristic is a bit long. If there was a way to describe each heuristic in one sentence, the set would be much easier to go through and understand."
- P2: "Two too many heuristics. If possible, a set of 10 works much better."
- P2: "Explanations are a bit too long. It requires extra work (cognitive load) for the users to understand Set C (Joyce et al.)."

6 Conclusion

HCI practitioners and researchers continue to use traditional usability evaluation methods to evaluate the usability of mobile applications. Yet, these methods were designed to evaluate desktop applications, and do not consider issues specific to mobile applications. In this work, the authors empirically investigate a claim from previous publications that one such method—Heuristic Evaluation—can be modified and consequently prove to be more effective in surfacing usability issues specific to mobile applications. Our study demonstrates that this is indeed the case. Additionally, participants felt most confident in using mobile application heuristics defined by the authors to evaluate usability of mobile applications in a professional context. However, the mobile application heuristics defined by the authors need further work; participants felt that the heuristics could be easier-to-use and to-learn, if they were reduced in number, yet were just as comprehensive, and had shorter descriptions.

This research is an important consideration for HCI practitioners and researchers responsible for the usability evaluations of mobile applications. Indeed, any teams responsible for the development of mobile applications can benefit from this work.

Acknowledgments The authors wish to thank the six HCI practitioners that participated in the Heuristic Evaluation and Evaluation of Heuristics.

References

1. Bernhaupt, R., Mihalic, K., Obrist, M.: Usability evaluation methods for mobile applications. In: Lumsden, J. (ed.) *Handbook of Research on User Interface Design and Evaluation for Mobile Technology*. IGI Global (2008)
2. Baharuddin, R., Singh, D., Razali, R.: Usability dimensions for mobile applications—a review. *Res. J. Appl. Sci. Eng. Technol.* **5**(6), 2225–2231 (2013)

3. Nielsen, J.: Smartphones: so many apps, so much time. Available Online: <http://www.nielsen.com/us/en/insights/news/2014/smartphones-so-many-apps-so-much-time.html>. (2014)
4. Coursaris, C.K., Kim, D.J.: A meta-analytical review of empirical mobile usability studies. *J. Usability Stud.* **6**(3), 117–171 (2011)
5. Joyce, G., Lilley, M.: Towards the development of usability heuristics for native smartphone mobile applications. In: *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience*, pp. 465–474 (2014)
6. Joyce, G., Lilley, M., Barker, T., Jefferies, A.: Smartphone application usability evaluation: the applicability of traditional heuristics. In: Marcus, A., (ed.) *Design, User Experience, and Usability: Users and Interactions*, (9187), pp. 541–550 (2015)
7. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 249–256 (1990)
8. Maguire, M.: Methods to support human-centred design. *Int. J. Hum Comput. Stud.* **55**(4), 587–634 (2001)
9. Jeffries, R., Desurvire, H.: Usability testing vs. heuristic evaluation: was there a contest? *ACM SIGCHI Bull.* **24**(4), 39–41 (1992)
10. Law, E.L.-C., Hvannberg, E.T.: Analysis of strategies for improving and estimating the effectiveness of heuristic evaluation. In: *Proceedings of the Third Nordic Conference on Human-Computer Interaction—NordICHI '04*, pp. 241–250 (2004)
11. Ketola, P., Roykkee, M.: Three facets of usability in mobile handsets. In: *Proceedings of CHI 2001 Workshop: Mobile Communications: Understanding Users, Adoption & Design* (2001)
12. Po, S., Howard, S., Vetere, F., Skov, M.B.: Heuristic evaluation and mobile usability: bridging the realism gap. *Mobile human-computer interaction—MobileHCI 2004. In: Lecture Notes in Computer Science*, vol. 3160/2004, pp. 591–592 (2004)
13. Weiss, S.: *Handheld usability*. Wiley (2002)
14. Ji, Y.G., Park, J.H., Lee, C., Yun, M.H.: A usability checklist for the usability evaluation of mobile phone user interface. *Int J. Hum Comput Interact.* **20**(3), 207–231 (2006)
15. Bertini, E., Gabrielli, S., Kimani, S.: Appropriating and assessing heuristics for mobile computing. In *AVI '06: Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 119–126 (2006)
16. Inostroza, R., Rusu, C., Roncagliolo, S., Jiménez, C., Rusu, V.: Usability heuristics for touchscreen-based mobile devices. In: *Proceedings of the 2012 Ninth International Conference on Information Technology—New Generations* (2012)
17. Nielsen, J.: Ten usability heuristics. Available at: http://www.useit.com/papers/heuristic/heuristic_list.html. (1994)
18. Nielsen, J.: Finding usability problems through heuristic evaluation. In: *Proceedings of the SIGCHI Conference on Human*, pp. 373–380 (1992)
19. Sauro, J.: Rating the severity of usability problems. Available Online: <http://www.measuringu.com/blog/rating-severity.php>. (2013)