Heuristic Evaluation of Mobile Usability: A Mapping Study

André de Lima Salgado and André Pimenta Freire

Federal University of Lavras
Computer Science Department
Campus Universitário, Caixa Postal 3037, Lavras, MG, Brazil
andre@bsi.ufla.br,
apfreire@dcc.ufla.br

Abstract. The mobile devices market has grown substantially, along with significant developments in mobile interactive technologies. Devices such as tablets, smartphones and others have increasingly become more popular and helped improve the way people interact and exchange information. The aim of this paper is to perform a systematic mapping of the literature regarding the use of heuristic evaluation methods applied on mobile applications. The aims of this research were twofold: analysing what are the most used sets of usability heuristics on usability evaluations of mobile devices, providing a common base to improve mobile design and usability evaluation; analysing details of how usability inspections of mobile applications have been conducted. The results show that different heuristics have been reported in research papers to evaluate usability of mobile devices. The study identified a total of 9 different heuristics sets means of the literature mapping. The traditional set of heuristics proposed by Nielsen and Molich was still the most used set of heuristics in heuristic usability evaluations of mobile devices, but the proposal of new specific heuristics for mobile interfaces has grown substantially.

Keywords: mobile usability, heuristic evaluation, usability inspection.

1 Introduction

There has been a substantial growth in the mobile devices market, along with significant developments in mobile interactive technologies [4] [27]. Devices such as tablets, smartphones and others have increasingly become more popular and helped improve the way people interact and exchange information. According to Monetate Ecommerce Quarterly [21], e-commerce businesses are seeking for ways to deliver relevant messages to mobile consumers because the traffic from smartphones has increased in a very fast pace.

Mobile devices have some specifics characteristics, they can be used while walking, within different weather situations and it can be easily used at different places [2].

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2014, LNCS 8512, pp. 178–188, 2014.

Within this context, the development of mobile applications has been an important segment in the software industry [4]. Given the key role played by usability as a quality attribute in software development [13], it has become extremely relevant to research appropriate methods to evaluate and improve the usability of mobile applications.

Although having become more ubiquitous, the usability of many mobile applications still remains a challenge to be addressed [4] [10]. According to Billi et al. [5], mobile devices have a number of limitations, such as small screen, limited input capabilities, limited computational resources, limited power supply (batteries), and large heterogeneity in models. Nayebi et al. [22] state that there is a need for more specific and systematic measurement methodologies for evaluating mobile usability.

Heuristic evaluation (HE), proposed by Nielsen and Molich [24], has been championed as a useful low-cost inspection method to professionally evaluate software usability. HE methods are extensively used to evaluate usability of conventional desktop software. This method requires that evaluators inspect some specific interface elements, being guided by pre-defined tasks and compare the interface with a list of heuristics aiming to find usability problems. However, considering the specificities and all differences between mobile and desktop technologies, adopting methods and acquired knowledge in the evaluation of desktop applications directly in the evaluation of mobile applications may not be necessarily straightforward [27].

The use of ordinary desktop-oriented heuristics in HE methods may not always be appropriate to evaluate all types of mobile interfaces, and may not embrace mobile-specific interface characteristics [7]. Heo et al. [10] argue that even with the large number of heuristics for usability evaluation of software systems, they are still not sufficient to appropriately reveal many mobile usability issues, and Bertini et. al. [2] afirm that specific heuristics are need because the traditional ones implicitly embody assumptions about static desktop location and use.

In order to help investigate the current state of the art to develop into these gaps, this paper aims to perform a systematic mapping of the literature regarding the use of heuristic evaluation methods applied on mobile applications. The goal of this research is to analyse what are the most used sets usability heuristics on usability evaluation of mobile devices, investigate how HE methods have been applied and to map the different heuristics that have been used on mobile usability evaluations to provide a common base to improve the design and evaluation of the usability of mobile applications.

The method employed was based on systematic reviews of the literature, by means of reviews and quantitative and qualitative analyses of data reported in literature, combining the results with meta-analytic techniques to increase the likelihood of discovering effects that smaller studies are not able to detect [17].

2 Literature Mapping Process

2.1 Description

Reviewing the literature to systematically collect studies related to a specific theme is a work of precision. The review process of the present study was made based on Kitchenham's [16] stages of a systematic review. As proposed by Kitchenham [16], this mapping process is based on three main phases: *Planning the Review*, *Conducting the Review* and *Reporting the Review*.

Within the *Planning the Review* phase, two stages are included:

1. Identification of the need for a review.
2. Development of a review protocol.

The *Conducting the Review* phase consists of five other stages:

1. Identification of research.
2. Selection of primary studies.
3. Study quality assessment.
4. Data extraction and monitoring.
5. Data synthesis.

The final phase, *Reporting the Review*, emcompasses the review of the whole mapping process, the analysis of all of the results and the publications of the work.

2.2 Mapping Review Protocol

The mapping review protocol includes all methods that are to be done by means of the systematic review. Without this protocol, there is a substantial chance that the mapping process be driven by researcher expectations [16]. The points of the Mapping review protocol are highlighted as following.

Objective: this mapping study has the main goal of finding what are the heuristics used to evaluate usability of mobile devices.

Research Question: "What are the main aspects of heuristic evaluations of mobile usability reported in literature?".

Source search method: sources will be searched by web engines using a search string.

Keywords: usability, heuristics, heuristic evaluation, mobile, android, iOS.

Sources List: publications of all types that were indexed by ACM Digital Library, IEE Xplore, SpringerLink and Science Direct.

Inclusion and exclusion criteria: publications considered in this mapping must be available in electronic format through the mentioned sources list and be written in english. Beside this, publications need to inform, at least, about the keywords mobile, heuristic and usability inside the summary.

Quality assessment criteria: publications need to inform about heuristics or heuristics sets used in usability evaluation of mobile devices. Each publication should also inform enough details about the procedure to perfom the inspection in order to verify whether it could be classified as a heuristic evaluation.

2.3 Conducting the Review

Ater the Mapping review protocol was completed, the mapping process was conducted by one research and supervised by another one (both authors of this paper).

The *identification of research* stage was conducted from October 16th, 2013, to November 7th, 2013. To execute all searches through chosen sources lists search engines, the following search string was used: "usability" AND ("heuristic evaluation" OR "heuristics") AND ("mobile" OR "android" OR "iOS"). A total of 2172 publications were identified through all source search engines listed. From these results, 26 are from ACM Digital Library, 10 from IEEE Xplore, 1335 from SpringerLink and 801 from Science Direct.

From the initial set of 2172 search results, the selection of primary studies (described previously) was performed to execute a first filter. This filter is a faster way to decrease the number of publications that do not overlap the mapping purposes. The *selection of primary studies* have occurred and 26 studies were accepted, 3 from the ACM Digital Library, 7 from IEEE Xplore, 13 from SpringerLink and 3 from Science Direct.

For a final filter, the study quality assessment criteria were applied. The goal of this phase was to analyse all results that provide evidence of using heuristics to mobile usability evaluation, in order to compose the final body of results of this systematic mapping. After the *study quality assessment* stage, 19 studies were included in the mapping process, being 3 from the ACM Digital Library, 5 from IEEE Xplore, 10 from SpringerLink and 1 from Science Direct.

All of the 19 studies were revised and the *data extraction and monitoring* phase was performed to collect information about the types of heuristics used in the evaluation and the process used to perform the heuristic evaluations in the studies. At this stage we used a database containing the full list of all identified heuristics, and other details about the method used to evaluate (such as whether the heuristic evaluations were performed according to the method proposed by Nielsen and Molich [24] or if it used some variation of the method) and the number of evaluators.

Finally, in the *data synthesis* of this mapping study, we analysed the main aspects observed in the literature about heuristic evaluation of mobile usability. This study reviewed a total of 2172 publications about the theme, and after using two filters, we obtained 19 results that have provided us a large list of heuristics used to evaluate usability of mobile devices. This list was reviewed by two coders, and after analysing heuristics with similar descriptions, a list of 29 distinct heuristics was obtained (see full list in the Appendix).

3 Results and Discussion

The results presented in this paper are the synthesis of the analysis of 19 publications that provide information about heuristic evaluation of mobile applications. We present in this section the results of the data analysis and summary of the findings. Section 3.1 presents the evolution of research works performing

heuristic evaluations of mobile usability, Section 3.2 presents a summary of the identification of distinct heuristics and heuristics sets used in the research works surveyed, and Section 3.3 presents an analysis of the methodological aspects of how the heuristic inspections were performed.

3.1 Evolution of Research Studies Using Heuristic Evaluation of Mobile Usability

The results including in this systematic mapping spanned a period ranging from 2004 to 2013. Figure 1 shows the number of research studies published in each year grouped by the database where they were sourced from.

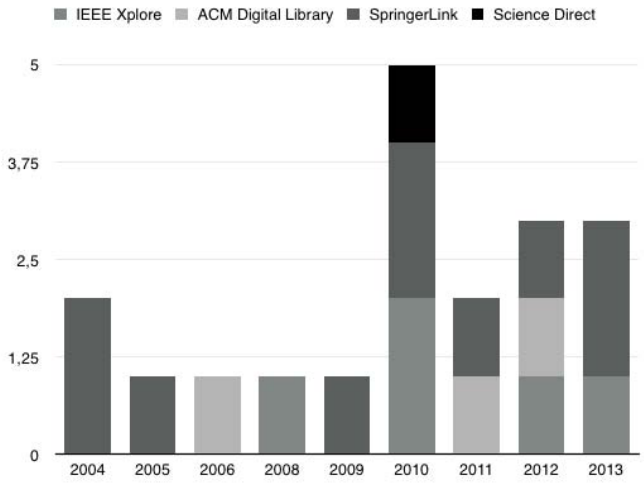

Fig. 1. Number of mapping results, divided by sources and year of publishing

It is possible to observe that the number of publications reporting the use of heuristic evaluation of mobile applications peaked in 2010. However, other research studies have still been published in the following years, which shows this is still a relevant research topic.

3.2 Sets of Usability Heuristics Used in Surveyed Studies

Identifying which heuristics are reported by research studies was an important goal of this systematic literature mapping. We aimed to collect and map the heuristics used in each paper and to verify whether significantly more new heuristics have emerged, or whether classic desktop-oriented heuristics are also applied in the evaluation of mobile applications.

In an initial analysis, we identified a total of 106 heuristics, belonging to 9 different sets. Following the initial identification, the authors coded similar

heuristics in order to identify a set of distinct heuristics. This resulted in a set of 29 distinct heuristics. The full list of the distinct heuristics identified is listed at the Appendix, containing the identification of each heuristic and which research studies considered it in the evaluation.

Table 1 summarises nine sets of heuristics identified as a set proposed by given authors. Table 1 also shows citations of research studies that mentioned using each set. It is possible to note that the heuristics purposed by Nielsen and Molich [24], one of the classic and most popular sets of heuristics to evaluate desktop software, were the most used to perform heuristic evaluation of mobile applications in the studies surveyed.

As broken down in the Appendix of this paper, heuristics worded exactly as defined by Nielsen and Molich's heuristics or heuristics very closely related to them were the most used in the surveyed studies. Other heuristics not cited by as many studies included domain-specific heuristics, especially those related to games.

Table 1. Heuristics sets and surveyed studies that reported using them

Heuristic set	**Referenced by studies**
Nielsen and Molich [24]	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14],
Heuristics for evaluating game usability [18]	[32], [29]
Mobile usability heuristics [3]	[3], [5], [6]
MATcH [27]	[27]
MMRGs [9]	[8]
Heuristics for Designing Mobile Applications [19]	[19]
UI Design Heuristics [20]	[20]
Heuristic Evaluation - A System Checklist [25]	[6]
Touchscreen-based mobile devices heuristics	[11]

Despite the differences in technologies, these results show that old principles traditionally used in HCI also apply to mobile usability, considering the specificities of those devices.

3.3 Methodology Used in Usability Inspections with Usability Heuristics

The analysis of the results also provided data about the evaluation method used, or suggested to future use, on each research study. According to the definition from Nielsen and Molich [24] of HE, we devided the research studies into those which followed the HE method more strictly according to the original definition

and others which somehow used a heuristics to perform other types of usability inspections (such as using heuristics in checklists and questionnaires). In this cases, the evaluation was taken asking evaluators to follow a static questionnaire, or a checklist, in order to find usability problems (which does not follow the definition of Nielsen and Molich [24] of HE). According to the results, 13 research studies followed Nielsen and Molich's [24] original method, and 6 studies showed heuristics used in different methods.

From the 13 studies that followed a stricter definition of heuristic evaluations as defined by Nielsen and Molich, 11 informed the number of evaluators involved on the HE. Of these 11 studies, one cited an HE performed by just two evaluators; four mentioned HE using three evaluators; five shoed HE performed by four evaluators and one reported an HE being performed by five evaluators. Figure 2 summarizes the number of studies that employed each number of evaluators. Most evaluations employed between 3 and 4 evaluators.

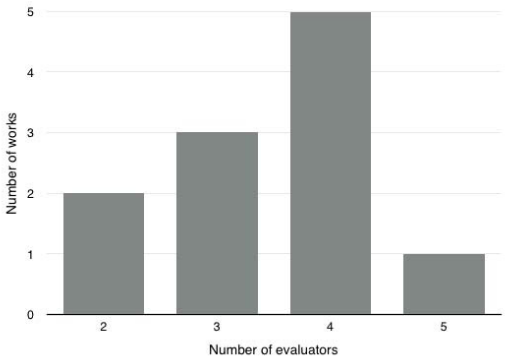

Fig. 2. Number of evaluators employed in HE studies

We also analyzed the severity scales used by different research studies. A total of 9 research studies categorized the severity of problems using the scale proposed by Nielsen, ranging from 1 - cosmetic, 2 - minor, 3 - major and 4 - catastrophic. Two studies used Likert-scales and one study categorized severities in low, medium or critical. The other studies did not provide details about the severity rate scale used.

4 Conclusion and Future Work

This work aimed to perform a systematic mapping of the literature regarding the use of heuristic evaluation methods applied on mobile applications. According to the results, this theme has been a relevant topic and that since 2010 the interest on it has been bigger than previous years.

The results also showed that a considerable number of heuristics sets has been used, a total of 9 heuristics sets were reported in the literature and the classic

heuristic set proposed Nielsen and Molich [24] was the most used set to evaluate usability of mobile applications. In most part of results, HE are taken following Nielsen and Molich [24] definition of HE method and the most common numbers of evaluators during HE of mobile usability are 3 and 4. Despite the large adoption of Nielsen and Molich [24] set of heuristics, some works proposes the use of heuristics along questionnaries and checklists, contradicting the Nielsen and Molich [24] definition of HE.

As future work, we intend to deepen the analysis of the heuristics surveyed in practice applying them in empirical studies with evaluators performing heuristic evaluations of mobile applications.

Acknowledgments. Thanks to ALCANCE Research Group, the Federal University of Lavras, CNPq, FAPEMIG and to MSc. Elisa Boari de Lima for all their support to this research.

References

1. Barricelli, B.R., Devis, Y., Abdelnour-Nocera, J., Wilson, J., Moore, J.: MANTRA: mobile anticoagulant therapy management. In: 2013 7th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), pp. 278–281. IEEE (2013)
2. Bertini, E., Catarci, T., Dix, A., Gabrielli, S., Kimani, S., Santucci, G.: Appropriating heuristic evaluation for mobile computing. Human-Computer Interaction and Innovation in Handheld, Mobile and Wearable Technologies 20 (2011)
3. Bertini, E., Gabrielli, S., Kimani, S.: Appropriating and assessing heuristics for mobile computing. In: Proceedings of the Working Conference on Advanced Visual Interfaces, pp. 119–126. ACM (2006)
4. Biel, B., Grill, T., Gruhn, V.: Exploring the benefits of the combination of a software architecture analysis and a usability evaluation of a mobile application. Journal of Systems and Software 83(11), 2031–2044 (2010)
5. Billi, M., Burzagli, L., Catarci, T., Santucci, G., Bertini, E., Gabbanini, F., Palchetti, E.: A unified methodology for the evaluation of accessibility and usability of mobile applications. Universal Access in the Information Society 9(4), 337–356 (2010)
6. Conde, A., et al.: An Intelligent Tutoring System Oriented to the Integration of People with Intellectual Disabilities. In: Demazeau, Y., Dignum, F., Corchado, J.M., Bajo, J., Corchuelo, R., Corchado, E., Fernández-Riverola, F., Julián, V.J., Pawlewski, P., Campbell, A. (eds.) Trends in PAAMS. AISC, vol. 71, pp. 639–647. Springer, Heidelberg (2010)
7. Fetaji, M., Fetaji, B.: Comparing developed MLUAT (mobile learning usability attribute testing) methodology with qualitative user testing method and heuristics evaluation. In: Proceedings of the 12th International Conference on Computer Systems and Technologies, pp. 516–523. ACM (2011)
8. Gielkens, C., Wetzel, R.: A framework for usability evaluation of mobile mixed reality games. In: Herrlich, M., Malaka, R., Masuch, M. (eds.) ICEC 2012. LNCS, vol. 7522, pp. 401–404. Springer, Heidelberg (2012)
9. Gielkens, C.: A framework for usability evaluation of mobile mixed reality games. Faculty of Science Theses (2012)

10. Heo, J., Ham, D.H., Park, S., Song, C., Yoon, W.C.: A framework for evaluating the usability of mobile phones based on multi-level, hierarchical model of usability factors. Interacting with Computers 21(4), 263–275 (2009)
11. Inostroza, R., Rusu, C., Roncagliolo, S., Jimnez, C., Rusu, V.: Usability Heuristics for Touchscreen-based Mobile Devices. In: 2012 Ninth International Conference on Information Technology: New Generations (ITNG), pp. 662–667. IEEE (2012)
12. Huang, C.H., Wang, C.M.: Usability Analysis in Gesture Operation of Interactive E-Books on Mobile Devices. In: Marcus, A. (ed.) Design, User Experience, and Usability, Pt I, HCII 2011. LNCS, vol. 6769, pp. 573–582. Springer, Heidelberg (2011)
13. ISO, I., FCD, I.: 25000, Software Engineering-Software Product Quality Requirements and Evaluation (SQuaRE)-Guide to SQuaRE. International Organization for Standardization, Geneva (2004)
14. Karahoca, A., Bayraktar, E., Tatoglu, E., Karahoca, D.: Information system design for a hospital emergency department: A usability analysis of software prototypes. Journal of Biomedical Informatics 43(2), 224–232 (2010)
15. Karampelas, P., Akoumianakis, D., Stephanidis, C.: Usability inspection of the WardInHand prototype. In: Universal Access in Health Telematics, pp. 197–208. Springer, Heidelberg (2005)
16. Kitchenham, B.: Procedures for performing systematic reviews. Keele, UK, Keele University 33 (2004)
17. Kitchenham, B.A., Charters, S.: Guidelines for performing systematic literature reviews in software engineering (2007)
18. Korhonen, H., Koivisto, E.M.: Playability heuristics for mobile games. In: Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services, pp. 9–16. ACM (2006)
19. Longoria, R.G., McGee, M., Nash, E.: Heuristics for Designing Mobile Applications. In: Designing Software for the Mobile Context, pp. 109–134. Springer, London (2004)
20. Marcus, A., Abromowitz, S., Abulkhair, M.F.: Heuristic Evaluation of iCalamityGuide Application. In: Marcus, A. (ed.) DUXU 2013, Part III. LNCS, vol. 8014, pp. 130–139. Springer, Heidelberg (2013)
21. Monetate: More Bad News for Email?. Ecommerce Quarterly (EQ3) (2013)
22. Nayebi, F., Desharnais, J.M., Abran, A.: The state of the art of mobile application usability evaluation. In: 2012 25th IEEE Canadian Conference on Electrical & Computer Engineering (CCECE), pp. 1–4. IEEE (2012)
23. Nielsen, J., Budiu, R.: Mobile usability. Pearson Education (2012)
24. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 249–256. ACM (1990)
25. Pierotti, D.: HE Checklist. Xerox Corporation, `http://www.stcsig.org/usability/topics/articles/he-checklist.html` (last accessed on February 7, 2014)
26. Po, S., Howard, S., Vetere, F., Skov, M.B.: Heuristic evaluation and mobile usability: Bridging the realism gap. In: Brewster, S., Dunlop, M.D. (eds.) Mobile HCI 2004. LNCS, vol. 3160, pp. 49–60. Springer, Heidelberg (2004)
27. Salazar, L.H.A., Lacerda, T., von Wangenheim, C.G., Barbalho, R.A.: Customization of usability heuristics for mobile phones. In: Companion Proceedings of the 11th Brazilian Symposium on Human Factors in Computing Systems, pp. 37–38. Brazilian Computer Society (2012) (in Portuguese)

28. Schönfelder, R., Schmalstieg, D.: Augmented reality for industrial building acceptance. In: Virtual Reality Conference, VR 2008, pp. 83–90. IEEE (2008)
29. Soomro, S., Wan Ahmad, W.F., Sulaiman, S.: Evaluation of Mobile Games Using Playability Heuristics. In: Zaman, H.B., Robinson, P., Olivier, P., Shih, T.K., Velastin, S. (eds.) IVIC 2013. LNCS, vol. 8237, pp. 264–274. Springer, Heidelberg (2013)
30. Varsaluoma, J.: Scenarios in the Heuristic Evaluation of Mobile Devices: Emphasizing the Context of Use. In: Kurosu, M. (ed.) HCD 2009. LNCS, vol. 5619, pp. 332–341. Springer, Heidelberg (2009)
31. Wahab, N.A., Osman, A., Ismail, M.H.: Engaging children to science subject: a heuristic evaluation of mobile learning prototype. In: 2010 Second International Conference on Computer Engineering and Applications (ICCEA), vol. 2, pp. 513–516. IEEE (2010)
32. Zaibon, S.B., Shiratuddin, N.: Heuristics evaluation strategy for mobile game-based learning. In: 2010 6th IEEE International Conference on Wireless, Mobile and Ubiquitous Technologies in Education (WMUTE), pp. 127–131. IEEE (2010)

Appendix:

#	**Heuristic**	**Referenced by studies**
1	Visibility of system status	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14], [3], [5], [27], [6], [32], [29], [20]
2	Match between system and the real world	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14], [3], [5], [6], [27], [20], [32], [29]
3	User control and freedom	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14], [27], [20], [6]
4	Consistency and standards	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14], [32], [29], [3], [5], [6], [27], [20], [9], [19]
5	Error prevention	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14], [20], [6], [8]
6	Recognition rather than recall	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14], [32], [29], [27], [8], [20], [6]
7	Flexibility and efficiency of use	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14], [32], [29], [3], [5], [6], [27], [20]
8	Aesthetic and minimalist design	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14], [3], [5], [27], [19], [20], [6], [8], [32], [29]
9	Help users recognize, diagnose, and recover from errors	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14], [32], [29], [8], [20], [6], [3], [5]
10	Help and documentation	[11], [31], [1], [28], [7], [26], [30], [12], [15], [14], [32], [29], [8], [20], [6]

#	**Heuristic**	**Referenced by studies**
11	Audio-visual representation supports the game	[32], [29], [8]
12	Pleasurable and Respectful Interaction with the User	[32], [29], [6]
13	Provide constant and appropriate feedback	[32], [29], [8], [20]
14	Ease of input, screen readability and glanceability	[3], [5], [6], [8], [19]
15	Aesthetic, privacy and social conventions	[3], [5], [6]
16	Physical interaction and ergonomics	[11], [27]
17	Information legibility and density	[27], [20]
18	Keep your navigation model simple	[27], [19]
19	Every round trip counts	[8], [19], [20]
20	There is a need	[19]
21	Provide unobstructed views that are appropriate for the users current situation	[8]
22	Real world navigation takes into account the type of game and is logical	[8]
23	Safeguard the players legal safety	[8]
24	Think modular	[19]
25	Allow for desktop based communication	[19]
26	Fight the hype	[19]
27	Direct manipulation/See and point	[20]
28	Modelessness	[20]
29	Perceived stability	[20]