User Experience Evaluations: Challenges for Newcomers

Cristian Rusu[1(✉)], Virginica Rusu[2], Silvana Roncagliolo[1], Juan Apablaza[1], and Virginia Zaraza Rusu[1]

[1] Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
{cristian.rusu, silvana}@ucv.cl, jjapablaza@gmail.com, rvzaraza90@hotmail.com
[2] Universidad de Playa Ancha, Valparaíso, Chile
virginica.rusu@upla.cl

Abstract. Human – Computer Interaction (HCI) should be a basic part of the formative process of all computer science professionals. Usability and User Experience (UX) were (re)defined by many authors and well recognized standards. UX is usually considered as an extension of usability. To move from usability to UX seems to be a tendency lately. Forming usability/UX evaluators is a challenging task. Practice is usually more appealing and persuasive than theory. The paper presents a study on the perception of (novice) evaluators over generic and specific usability heuristics.

Keywords: Usability · User experience · Usability evaluation · Heuristic evaluation · Usability heuristics

1 Introduction

The Joint Task Force on Computing Curricula of Association for Computing Machinery (ACM) and IEEE Computer Society establishes Human – Computer Interaction (HCI) as part of the Body of Knowledge in their Computer Science (CS) curricula proposal (CS2013) [1].

Usability was (re)defined by many authors. Usability definitions were also provided by well recognized standards. One of the best known and widely used definitions is the one proposed by ISO 9241: the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [2].

The UX concept was also referred by ISO 9241: a person's perceptions and responses that result from the use or anticipated use of a product, system or service [2]. Some authors consider UX as an extension of the usability concept. Others use the terms usability and UX indistinctly. The UX concept is still under review. The "User Experience White Paper" aims to "bring clarity to the concept" [3].

CS2013 explicitly includes usability as a compulsory core HCI topic. Usability is also recommended as elective topic. User Experience (UX) is not explicitly incorporated as a core HCI topic; however it is implicitly considered in other core and elective topics. It seems that the usability concept is widely accepted not only by the HCI

A. Marcus (Ed.): DUXU 2015, Part I, LNCS 9186, pp. 237–246, 2015.
DOI: 10.1007/978-3-319-20886-2_23

community, but also by the CS community in general. The UX concept is not yet commonly endorsed by the CS community. To move from usability to UX seems to be a tendency lately. Even the former "Usability Professionals Association" (UPA) redefined itself as "User Experience Professionals Association" (UXPA).

Including HCI in the CS curricula is still a challenge in most Latin American (LA) countries [4]. However, when HCI is present usability seems to be a major topic in both teaching and researching.

Forming usability/UX evaluators is a challenging task. The paper presents an empiric study on the perception of (novice) evaluators over generic and specific usability heuristics. Section 2 examines the concept of usability and UX. Section 3 highlights the heuristic evaluation as a fundamental assessment method for both usability and UX, examining the importance of the set of heuristics that are employed. Section 4 presents the results of experiments that were recently made, using both generic and specific usability heuristics. Section 5 points out conclusions and future work.

2 Usability and User Experience

Over more than three decades usability was (re)defined by many authors. Usability definitions were also provided by well recognized standards. Lewis points out that there is still no clear and generally accepted usability definition, as its complex nature is hard to describe in one definition [5].

One of the best known and widely accepted usability definitions was proposed by ISO 9241 standard: "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [6]. Updated ISO standards still refer to the ISO 9241 usability definition. ISO/IEC 25010 defines usability as the "degree to which a product or system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [7]. It considers usability as a subset of quality in use consisting of "effectiveness", "efficiency" and "satisfaction".

Nielsen and Loranger define usability as a quality attribute relating to how easy something is to use [8]. More specifically, how quickly people can learn to use it (learnability), how efficient they are using it (efficiency), how memorable it is (memorability), how error-prone it is (errors), and how much users like using it (satisfaction). Sharp, Rogers and Preece affirm that usability is generally regarding as ensuring that interactive products are easy to learn, effective to use, and enjoyable [9]. They denote six "usability goals": effectiveness, efficiency, safety, utility, learnability and memorability. Usability.gov states that usability is about effectiveness, efficiency and the overall satisfaction of the user, a combination of factors including: intuitive design, ease of learning, efficiency of use, memorability, error frequency and severity, and subjective satisfaction [10].

Regardless they are called "attributes", "factors" or "goals", usability dimensions are recurrent in all definitions. They are also referred in ISO standards. New interaction paradigms, new technologies and new kind of software systems are compelling

arguments for reviewing the usability concept, characteristics, methods, metrics and methodologies [11, 12].

ISO 9241-210 standard defines UX as a "person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service" [2]. It considers that UX "includes all the users' emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and accomplishments that occur before, during and after use".

Sharp, Rogers and Preece point out that one cannot design a user experience, but only design for a user experience; one cannot design a sensual experience, but only create the design features that can evoke it [9]. They enumerate a broad range of UX positive and negative "qualities". Kuniavsky admits that defining UX is difficult, since it can extend to nearly everything in someone's interaction with a product [13]. UXPA.org defines UX as every aspect of the user's interaction with a product, service, or company that make up the user's perceptions of the whole [14].

The UX concept is still under review. The "User Experience White Paper" aims to "bring clarity to the UX concept" [3]. Rather than intending to give a unique UX definition, the document mentions the wide collection of definitions available at All-aboutux.org [15].

Some authors consider UX as an extension of the usability concept. Others use the terms usability and UX indistinctly.

The ISO 9241-210 standard sustains that usability, when interpreted from the perspective of the users' personal goals, can include the kind of perceptual and emotional aspects typically associated with user experience [2]. Usability criteria can be used to assess aspects of user experience.

Usability.gov refers to usability as the quality of a user's experience when interacting with products or systems [10]. UXPA.org makes a direct link between UX and usability, through the human-centered design (HCD) process [14].

Lewis considers user-centered design (UCD) and UX as usability extensions; in his opinion UCD subsumed usability engineering (and ergonomics and human factors engineering), and UX has subsumed UCD [5]. He also points out that in the (near) future UX will probably become part of a larger customer experience effort, as a result of the growing emphasis on service design and the emergence of the service science as discipline. Lewis highlights usability as a stable component throughout the transformations from usability engineering to UCD to UX.

The "User Experience White Paper" considers that UX is not the same as usability, although usability, as perceived by the user, is typically an aspect contributing to the overall UX [3].

Sharp, Rogers and Preece make a distinction between usability goals and UX goals [9]. They consider usability goals concerned with meting specific usability criteria (related to effectiveness, efficiency and the overall satisfaction of the use), and UX goals concerned with explaining the nature of the UX itself. Overall, they consider usability goals more objective than UX goals.

Mitchell acknowledges that many terms are similar to usability [16]. He even considers that for practical purposes they are the same as usability: usability testing, human factors engineering, customer experience management, ergonomics, UCD, human factors, user – friendly design.

To move from usability to UX seems to be a tendency lately. Even the former "Usability Professionals Association" (UPA) redefined itself as "User Experience Professionals Association" (UXPA) [14].

3 Usability and User Experience Evaluation

Lewis highlights two major conceptions of usability [5]:

- Summative, focused on metrics (i.e., "measurement-based usability"),
- Formative, focused on usability problems detection and associated design solutions (i.e., "diagnostic usability").

The concept of summative usability led to ISO usability standard, emphasizing on three key factors: effectiveness, efficiency and satisfaction. Lewis highlights the resemblance of metrics associated to the three key factors to methods and metrics of experimental psychology, instantiated in human factors engineering.

The concept of formative usability focuses on the iterative design process (design – test – redesign). It led to the development of several usability evaluation methods, essentially classified as:

- Empirical usability testing, based on users' participation [17],
- Inspection methods, based on experts' judgment [18].

UX is generally considered an extension of usability; therefore usability evaluation methods may also be applied in order to assess UX. A broad collection of UX evaluation methods is provided at http://www.allaboutux.org/ [15].

Heuristic evaluation is one of the most common usability assessment methods. It involves the participation of usability specialists analyzing every interactive element and dialog following a set of established usability design principles called heuristics [19]. A heuristic evaluation is usually performed by 3 to 5 evaluators.

When selecting the set of heuristics to be used, there are (mainly) two alternatives: choosing generic heuristics or specific heuristics. Specific heuristics may become hard to understand and hard to apply but they can detect many (relevant) usability issues related to the application area. Generic heuristics are easy to understand and to apply, but they can miss specific usability issues [11].

Our research work over the last decade focused mainly on usability/UX and related topics. We came to the conclusion that the traditional usability engineering concepts and evaluation methods should be re-examined. There is a need for new evaluation methods or at least for the use of traditional evaluations in novel ways. Frameworks of usability evaluation, including appropriate methods or combination of methods should be established, in order to get more effective and efficient evaluations on new interaction paradigms. We proposed specific usability heuristics and associated checklists for transactional web applications [20], touchscreen-based mobile applications [21], grid computing applications [22], interactive digital television [23], and virtual worlds [24]. We also developed a cultural – oriented usability heuristics proposal [25]. The experience of developing specific usability heuristics led to a methodology proposal [11, 26].

4 Challenges When Forming Usability Professionals

Including HCI in the CS curricula and forming usability/UX professionals is still a challenge in most LA countries [4]. An appealing way to introduce HCI at all CS curricula levels is by systematically including usability/UX practices [27]. SIGCHI acknowledges the importance of getting down HCI to the practical work [28]. We believe that a strong relationship between HCI theory, research and practice is particularly important in countries were HCI communities are not yet well established.

Teaching HCI in Chile for more than a decade was a challenging intercultural, interdisciplinary, cross-field but very rewarding experience [29]. As practice is usually more appealing and persuasive than the theory, we gradually increased the weight of practical activities, and we came to focus more and more on teaching the students how to put the HCI theory into practice. Forming usability/UX evaluators is a challenging task.

Heuristic evaluations and usability tests are compulsory practice for all our students, at undergraduate and graduate level. As standard practice, at least one heuristic evaluation is performed based on Nielsen's set of 10 usability heuristics [19]. Usually a heuristic evaluation based on domain-specific usability heuristic is also performed. After each heuristic evaluation a standard questionnaire is applied, giving us an interesting and very useful feedback. Some results were previously published [30, 31].

An experiment was made in 2014, involving 54 CS students:

- 25 undergraduate students from Pontificia Universidad Católica de Valparaíso, Chile, 12 of them having some previous experience in heuristic evaluations,
- 29 graduate students from Pontificia Universidad Católica del Perú, all of them novice evaluators.

All participants were asked to perform a heuristic evaluation over a major airline transactional website (www.lan.com), using Nielsen's 10 usability heuristics. Later on a survey was conducted in order to evaluate their perception over Nielsen's heuristics, concerning 4 dimensions: D1 - Utility, D2 - Clarity, D3 - Ease of use, D4 - Necessity of additional checklist. All dimensions were evaluated using a 5 points Likert scale. As samples are independent, observations are ordinal, and no assumption of normality can be made, results were analyzed using nonparametric statistics tests.

A Mann-Whitney U test was performed to check the hypothesis:

- H0: there are no significant differences between evaluators with and without previous experience,
- H1: there are significant differences between evaluators with and without previous experience.

As decision rule was used $p \leq 0.05$. As Table 1 shows, there are no significant differences between the two groups of evaluators (with/without previous experience), excepting the dimension D4 - Necessity of additional checklist. A preliminary explanation could be that novice evaluators do not really understand the purpose of a checklist that complements a set of usability heuristics.

Table 1. Mann-Whitnew U test for the perception of Nielsen's heuristics

	D1: Utility	D2: Clarity	D3: Ease of use	D4: Necessity of additional checklist
p	0.1353	0.1921	0.0908	0.0359

A Spearman ρ test was performed to check the hypothesis:

- H0: p = 0, the dimensions Dm and Dn are independent,
- H1: p ≠ 0, the dimensions Dm and Dn are dependent.

As decision rule was used p ≤ 0.05. Results show that:

- All dimensions are independent in the case of evaluators with previous experience;
- There are some weak to moderate dependences in the case of novice evaluators (Table 2);
- There are also some weak to moderate dependences when all evaluators are considered (Table 3).

Table 2. Spearman ρ test for novice evaluators (Nielsen's heuristics)

	D1: Utility	D2: Clarity	D3: Ease of use	D4: Necessity of additional checklist
D1	1	0.55	Independent	Independent
D2		1	0.46	0.35
D3			1	Independent
D4				1

Table 3. Spearman ρ test for all evaluators (Nielsen's heuristics)

	D1: Utility	D2: Clarity	D3: Ease of use	D4: Necessity of additional checklist
D1	1	0.53	0.27	Independent
D2		1	0.43	Independent
D3			1	Independent
D4				1

A second experiment was made, involving the 25 undergraduate students from Pontificia Universidad Católica de Valparaíso, Chile. A month after the first experiment they were asked to perform a new heuristic evaluation over the same transactional website (www.lan.com), this time using a set of 14 specific usability heuristics for transactional websites (TW):

- TW1: Visibility of system status,
- TW2: Feedback on transactions' state,
- TW3: Dependability and anticipated functionality,
- TW4: Security and speed of transactions,
- TW5: Match between system and user world,
- TW6: User control and freedom,

- TW7: Consistency in system design,
- TW8: Use of web standards and symbols,
- TW9: Error prevention,
- TW10: Minimize user's memory load,
- TW11: Flexibility and efficiency of use,
- TW12: Aesthetic and minimalist design,
- TW13: Help for error recognition and recovery,
- TW14: Help and documentation.

The set of TW usability heuristics was developed in 3 iterations [11]. It was validated through several experiments and case studies [20, 31].

The website used as case study did not experienced any changes between the two experiments. A new survey was conducted in order to assess evaluators' perception over TW heuristics, concerning the same 4 dimensions and using the same scale as in the case of Nielsen's heuristics. Results were analyzed in similar ways.

As Table 4 shows, there are no significant differences between the two groups of evaluators (with/without previous experience). Apparently TW heuristics are clear enough; there's no significant need for additional checklist.

Table 4. Mann-Whitnew U test for the perception of TW heuristics

	D1: Utility	D2: Clarity	D3: Ease of use	D4: Necessity of additional checklist
p	0.7441	0.1708	0.6020	0.7692

The Spearman ρ test shows that:

- The major dependences in the case of evaluators with previous experience occurs between dimensions D2 – D1 and D2 – D3 (Table 5); if heuristics are perceived as clear (easy to understand), they are also perceived as useful and easy to use;
- There is a very high (and unexpected) opposite dependency between D3 and D1 in the case of novice evaluators (Table 6); a possible explanation would be their lack of experiences, however further analysis is required!
- There are some moderate dependences when all evaluators are considered (Table 7).

Additional (descriptive) statistics offered important feedback in order to refine the set of TW heuristics. Based on the lowest score, heuristics TW1, TW2, TW5, TW6, TW7, TW11, and TW12 definitions need further review.

Table 5. Spearman ρ test for evaluators with previous experience (TW heuristics)

	D1: Utility	D2: Clarity	D3: Ease of use	D4: Necessity of additional checklist
D1	1	0.68	0.49	Independent
D2		1	0.66	Independent
D3			1	Independent
D4				1

Table 6. Spearman ρ test for novice evaluators (TW heuristics)

	D1: Utility	D2: Clarity	D3: Ease of use	D4: Necessity of additional checklist
D1	1	Independent	-0.90	Independent
D2		1	Independent	Independent
D3			1	Independent
D4				1

Table 7. Spearman ρ test for all evaluators (TW heuristics)

	D1: Utility	D2: Clarity	D3: Ease of use	D4: Necessity of additional checklist
D1	1	0.54	Independent	Independent
D2		1	0.59	0.43
D3			1	Independent
D4				1

5 Conclusions

Over more than three decades usability was (re)defined by many authors. UX is generally considered an extension of usability. Some authors use the terms usability and UX indistinctly. New interaction paradigms, new technologies and new kind of software systems are compelling arguments for reviewing the usability and UX concepts, characteristics, methods, metrics and methodologies.

Including HCI in the CS curricula and forming usability/UX professionals is still a challenge in most LA countries. SIGCHI acknowledges the importance of getting down HCI to the practical work; an appealing way to introduce HCI at all CS curricula levels is by systematically including usability/UX practices. A strong relationship between HCI theory, research and practice is particularly important in countries were HCI communities are not yet well established.

Forming usability/UX evaluators is a challenging task. A study on the perception of (novice) evaluators over generic and specific usability heuristics was conducted. There are no significant differences between the groups of evaluators with and without previous experience, except the perceived necessity of additional checklist (when working with general heuristics). When occur, dependencies between the four surveyed dimensions are somehow expected. The only unexpected (very high, opposite) dependency occurs between dimensions “Ease of use” and “Utility”, in the case of novice evaluators (when working with specific heuristics for transactional websites).

The study offered an important feedback for both teaching and research. More experiments are necessary in order to validate preliminary conclusions. Quantitative analyze will be complemented with qualitative data, collected through surveys and interviews.

References

1. CS2013: Computer Science Curricula 2013. Curriculum Guidelines for Undergraduate Degree Programs in Computer Science (Final Report). ACM/IEEE-CS Joint Task Force on Computing Curricula, ACM/IEEE Computer Society (2013)
2. ISO 9241-210: Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. International Organization for Standardization, Geneva (2010)
3. Roto, V., Law, E., Vermeeren, A., Hoonhout, J.: User Experience White Paper. Bringing Clarity to the Concept of User Experience. http://www.allaboutux.org/uxwhitepaper. Accessed 20 Nov 2014
4. Collazos, C., Granollers, T., Rusu, C.: A survey of human-computer interaction into the computer science curricula in Iberoamerica. In: 8th International Conference on Information Technology: New Generations (ITNG 2011), pp. 151–156. IEEE Computer Society Press (2011)
5. Lewis, J.: Usability: lessons learned… and yet to be learned. Int. J. Hum.-Comput. Interact. **30**(9), 663–684 (2014)
6. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability. International Organization for Standardization, Geneva (1998)
7. ISO/IEC 25010: Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models. International Organization for Standardization, Geneva (2011)
8. Nielsen, J., Loranger, H.: Prioritizing Web Usability. New Riders, Berkley (2006)
9. Sharp, H., Rogers, Y., Preece, J.: Interaction Design: Beyond Human-Computer Interaction. Wiley, Chichester (2007)
10. Usability.gov: Improving the User Experience. U.S. Department of Health & Human Services. http://www.usability.gov/. Accessed 20 Nov 2014
11. Rusu, C., Roncagliolo, S., Rusu, V., Collazos C.: A methodology to establish usability heuristics. In: The Fourth International Conference on Advances in Computer-Human Interactions (ACHI 2011), pp. 59–62. IARIA (2011)
12. Wiberg, C., Jegers, K., Desurvire, H.: How applicable is your evaluation methods – really? In: The Second International Conference on Advances in Computer-Human Interactions (ACHI 2009), pp. 324–328. IEEE Computer Society Press (2009)
13. Kuniavsky, M.: Observing the User Experience: A Practitioner's Guide to User Research. Morgan Kaufmann, San Francisco (2003)
14. UXPA.org: User Experience Professionals Association. http://uxpa.org/. Accessed 20 Nov 2014
15. Allaboutux.org: All About UX. http://www.allaboutux.org/. Accessed 20 Nov 2014
16. Mitchell, P.: A Step-by-Step Guide to Usability Testing. iUniverse, New York (2007)
17. Dumas, J., Fox, J.: Usability testing: current practice and future directions. In: Sears, A., Jacko, J. (eds.) The Human – Computer Interaction Handbook: Fundamentals. Evolving Technologies and Emerging Applications, pp. 1129–1149. Taylor & Francis, New York (2008)
18. Cockton, G., Woolrych, A., Lavery, D.: Inspection – based evaluations. In: Sears, A., Jacko, J. (eds.) The Human – Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, pp. 1171–1189. Taylor & Francis, New York (2008)
19. Nielsen, J., Mack, R.L.: Usability Inspection Methods. Wiley, New York (1994)
20. Quiñones, D., Rusu, C., Roncagliolo, S.: Redefining usability heuristics for transactional web applications. In: 11th International Conference on Information Technology: New Generations (ITNG 2014), pp. 260–265. IEEE Computer Society Press (2014)

21. Inostroza, R., Rusu, C., Roncagliolo, S., Rusu, V.: Usability heuristics for touchscreen-based mobile devices: update. In: First Chilean Conference on Human - Computer Interaction (ChileCHI 2013), pp. 24–29. ACM International Conference Proceeding Series (2013)
22. Roncagliolo, S., Rusu, V., Rusu, C., Tapia, G., Hayvar, D., Gorgan, D.: Grid computing usability heuristics in practice. In: 8th International Conference on Information Technology: New Generations (ITNG 2011), pp. 145–150. IEEE Computer Society Press (2011)
23. Solano, A., Rusu, C., Collazos, C., Arciniegas, J.: Evaluating interactive digital television applications through usability heuristics. Ingeniare **21**(1), 16–29 (2013)
24. Rusu, C., Muñoz, R., Roncagliolo, S., Rudloff, S., Rusu, V., Figueroa, A.: Usability heuristics for virtual worlds. In: The Third International Conference on Advances in Future Internet (AFIN 2011), pp. 16–19. IARIA (2011)
25. Diaz, J., Rusu, C., Pow-Sang, J., Roncagliolo, S.: A cultural – oriented usability heuristics proposal. In: First Chilean Conference on Human - Computer Interaction (ChileCHI 2013), pp. 82–87. ACM International Conference Proceeding Series (2013)
26. Jimenez, C., Rusu, C., Roncagliolo, S., Inostroza, R., Rusu, V.: Evaluating a methodology to establish usability heuristics. In: 31st International Conference of the Chilean Computer Science Society (SCCC 2012), pp. 51–59. IEEE Computer Society Press (2013)
27. Rusu, C., Rusu, V., Roncagliolo, S.: Usability practice: the appealing way to HCI. In: The First International Conference on Advances in Computer-Human Interactions (ACHI 2008), pp. 265–270. IEEE Computer Society Press (2008)
28. ACM SIGCHI: ACM SIGCHI Curricula for Human-Computer Interaction. http://old.sigchi.org/cdg/cdg2.html#2_1. Accessed 12 Nov 2014
29. Rusu, C., Rusu, V.: Teaching HCI: a challenging intercultural, interdisciplinary, cross-field experience. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, pp. 344–354. Springer, Heidelberg (2007)
30. Jimenez, C., Rusu, C., Rusu, V., Roncagliolo, S., Inostroza, R.: Formal specification of usability heuristics: how convenient it is? In: 2nd International Workshop on Evidential Assessment of Software Technologies (EAST 2012), pp. 55–60. ACM International Conference Proceeding Series, (2012)
31. Paz, F., Villanueva, D., Rusu, C., Roncagliolo, S., Pow-Sang, J.A.: Experimental evaluation of usability heuristics. In: 10th International Conference on Information Technology: New Generations (ITNG 2013), pp. 119–126. IEEE Computer Society Press (2013)