Computer Standards & Interfaces 66 (2019) 103345

Contents lists available at ScienceDirect

Computer Standards & Interfaces

journal homepage: www.elsevier.com/locate/csi

Applying a methodology to develop user eXperience heuristics

Daniela Quiñones*, Cristian Rusu

Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile

ARTICLE INFO

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

ABSTRACT

A heuristic evaluation method allows the evaluation of the usability of application domains. To evaluate applications that have specific domain features, researchers can use sets of specific usability heuristics in addition to the well-known (usually Nielsen's) heuristics. Heuristics can also focus on the User eXperience (UX) aspects other than the usability. In a previous work, we proposed a formal methodology for establishing usability/UX heuristics. The methodology has 8 stages including activities to formulate, specify, validate and refine a new set of heuristics for a specific application domain. The methodology was validated through expert opinion and several case studies. Although when specifying the methodology, we explained each of its stages in detail, some activities can be difficult to perform without a guide that helps the researcher determine how the stages should be carried out. This article presents a detailed explanation regarding how to apply each stage of the methodology to create a new set of heuristics for a specific domain. Additionally, this paper explains how to iterate the methodology's stages and when to stop the process of developing new heuristics.

1. Introduction

Developing new heuristics as evaluation instruments has become a crucial activity to evaluate not only usability and User eXperience (UX) but also the specific features of different types of application domains. It is important to evaluate the application domain while considering all its components, features and the usability/UX attributes (or factors) that are critical in a specific application domain.

A widely used inspection method to evaluate usability, and the UX is a type of heuristic evaluation [1]. In this method, experts inspect a product/system/service to identify usability/UX problems and "measure" the usability/UX degree according to usability/UX principles or usability/UX heuristics [2]. Heuristics are rules or "guidelines" used to perform a heuristic evaluation [3]. Heuristic evaluation is a useful method to evaluate the usability/UX and detect potential problems. However, it may be necessary to use specific heuristics for the evaluation. Because each application domain may have unique and specific features, specific heuristics are recommended including elements that evaluate both usability/UX attributes/factors and their specific features [3].

There are methodologies that support the process of developing heuristics [4–10]; however, there has been no clear protocol for heuristic validation. We formalized the entire process of developing heuristics and proposed a formal methodology for establishing usability/user experience heuristics [11], including detailed activities to formulate, specify, validate and refine a new set of heuristics. The methodology has 8 stages to develop heuristics, which considers the features of a specific application domain (product/system), the usability and the UX. The methodology was validated in several iterations [11]. Based on the results obtained, we concluded that the methodology is perceived as useful and that researchers will use it in the future [11]. In addition, we also concluded that the methodology is effective since it has been used to create new sets of heuristics (a final "product" is obtained) [11].

Although when specifying the methodology, we explained in detail each of its stages, some activities can be difficult to perform without a guide that helps the researcher determine how these stages should be developed. This article presents a detailed explanation regarding how to apply each stage of the methodology to create a new set of usability/UX heuristics for a specific application domain. The article also explains how to iterate the methodology's stages and when to stop the process of developing new heuristics. We aim to help the researcher during the heuristic development process using this methodology. To explain the process of developing heuristics, the development of a set of UX heuristics for national park websites is used as a case study [12]. In addition, several other case studies are analyzed to complement the explanation.

This document is organized as follows: Section 2 explores the theoretical background; Section 3 briefly presents the methodology for developing usability/UX heuristics; Section 4 explains in detail how to apply each methodology stage and how to iterate to create a new set of

* Corresponding author.
E-mail addresses: daniela.quinones@pucv.cl (D. Quiñones), cristian.rusu@pucv.cl (C. Rusu).

https://doi.org/10.1016/j.csi.2019.04.004
Received 29 January 2019; Received in revised form 20 March 2019; Accepted 6 April 2019
Available online 08 April 2019
0920-5489/ © 2019 Elsevier B.V. All rights reserved.

heuristics; and Section 5 presents the conclusions and future studies.

2. Theoretical background

Usability is usually defined as the "capability of being used". The complex and evolving nature of usability is difficult to describe in a unique definition. Usability refers to ease of use and the way users can perform their tasks. UX is a broader concept that includes usability. UX covers all aspects of an individual's interaction with a product, system and/or service. There are more than 86 methods to evaluate usability and UX [13]. In general, evaluation methods can be classified into inspections, i.e., where expert evaluators participate; and tests, where users participate. One of the most widely used inspection methods is heuristic evaluation [1]. The concepts of usability, heuristic evaluation, and user experience are briefly presented below.

2.1. Usability

Usability is more than "ease of use". Usability involves a user interacting with a product, system or service; the goals that he/she expects to fulfill through the interaction; and the context of use, in addition to the efficiency, effectiveness and satisfaction associated with the interaction. The ISO 9241-11 standard defines usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [14]. Usability is a complex issue [15]. Each user (or group of users) may have different goals for the same product, system or service, depending on their needs. In addition, the interaction of each user can be different depending on the context of use. However, independent of the user, its goals and the context of use, the product, system or service must work in an efficient and effective manner, generating positive satisfaction for the user.

2.2. User eXperience

(UX) is a broader concept that includes usability; it involves not only the user's satisfaction, i.e., how pleasant it is to use a product, system or service, but also their emotions and perceptions during the interaction. The ISO 9241-210 standard defines the UX as "person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service" [14]. In other words, the UX includes all the beliefs, preferences, behaviors, physical and psychological responses that occur before, during and after use [14]. Currently, it is not enough to design products that are useful and easy to use. Users also look for products that create pleasant experiences, generating a bond beyond the functional.

Several authors have proposed aspects, attributes or factors that define the UX. These factors encompass the different dimensions and/or features that are part of UX. Morville proposes seven UX facets or factors to illustrate the UX: useful, usable, desirable, findable, accessible, credible, and valuable [16]. The facets are presented as a honeycomb diagram to help people understand the need to define priorities. Is it more important for a website to be desirable or accessible? How regarding usable or credible? The answer depends on the balance among the context of use, content and users [16].

Revang has suggested several factors that influence the UX [17]. The author states that the user experience is a series of phases and that it is important to focus on positivity in the six main factors: usability, usefulness, findability, accessibility, credibility, and desirability. In turn, each factor is subdivided into other factors with a total of 30 factors. Arhippainen and Tähti have created a model that shows five elements that influence the UX [18]. All these factors influence the experience that user product interaction evokes: social factors (such as time pressure), cultural factors (including language and habits), product factors (such as functions, usability, mobility, among others), user factors (e.g., emotions, skills, prior experiences, motivation, among others), and context of use (such as time and place). Garret has proposed several elements of the UX divided into five planes [19]: surface, skeleton, structure, scope, and strategy. These five planes provide a conceptual framework for discussing UX problems and the tools that one can use to solve them. These planes are not attributes but are presented as a guide for UX design.

It is important to consider each of the attributes/factors/facets of the UX when designing new products, systems or services. Likewise, to evaluate the UX, it is important to apply the most appropriate method (s) to evaluate all the factors or facets. Evaluating the UX is a challenge. A single evaluation method may not be sufficient to evaluate it completely. It is advisable to use both inspection methods and user tests to evaluate each UX aspect/attribute/factor. In this sense, the heuristic evaluation method allows evaluating the UX using a specific set of heuristics. However, since certain attributes of the UX are subjective (such as "desirable" and "valuable" [16]), and it can be difficult to fully evaluate the UX through a set of heuristics. A heuristic evaluation allows us to effectively evaluate certain attributes of the UX (useful, usable, findable, accessible, and credible [16]), but it is important to complement the evaluation using other methods to evaluate the UX completely.

To perform an effective evaluation of the UX and its factors, it is essential to use a set of specific heuristics. For this, it is possible to create a new set of heuristics if there is not one that evaluates the UX facets. Based on the results that have been presented in systematic literature reviews [20,21], apparently, there are no specific sets of heuristics to evaluate the UX and/or to evaluate certain attributes of the UX. However, it is critical to have sets of specific heuristics to correctly evaluate the UX through heuristic evaluation.

2.3. Heuristic evaluation

Heuristic evaluation is an inspection method used to evaluate the usability of a product interface. In addition, it is possible to evaluate certain UX aspects or attributes depending on the set of heuristics used. In this method, a set of expert evaluators inspect a product (interface) based on heuristics to identify usability/UX problems [2]. The problems are classified (associated with the heuristics used) and rated (in terms of severity, frequency and criticality).

Heuristics are "rules of thumb" or "usability guidelines" used in a heuristic evaluation as an instrument for detecting usability problems [3]. Nielsen's 10 heuristics [3] is one of the most commonly used sets of heuristics to evaluate systems or applications in general. However, since Nielsen's heuristics evaluate only general aspects or features of the systems, several sets of new usability heuristics have been developed to evaluate specific features of specific application domains. Based on the results presented in [20,21], more than 80 new sets of usability heuristics have been developed for specific domains.

3. The methodology

We proposed a formal methodology for developing usability/user experience heuristics for specific application domains [11]. The methodology has 8 stages that can be iteratively applied to refine and improve the new set of proposed heuristics. We validated the methodology in two iterations through expert opinions. The validation results are detailed in [11]. The methodology has already been applied to develop several new sets of heuristics for specific applications domains, such as online travel agency applications, websites, social networks, and videogames. All of these case studies proved the effectiveness of the methodology.

The methodology can be used to create heuristics and/or checklists that evaluate the usability or other UX aspects, such as playability, learnability, and communicability. This approach can also be used to create heuristic instruments for evaluating quality attributes other than usability (such as security and adaptability). Fig. 1 shows the 8

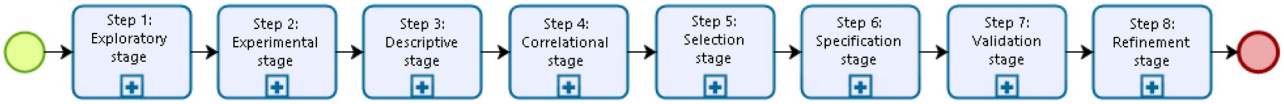

Fig. 1. Stages of the methodology to develop usability/user experience heuristics.

methodology stages. Although Fig. 1 sequentially shows the stages of the methodology, the development of heuristics may be performed iteratively. In specific situations, some stages may be optional, some stages overlap, and/or a stage may stop, and one can return to an earlier stage.

For each stage of the methodology, we specified (1) a definition, (2) the inputs that are needed to start the stage, (3) the activities that are performed in the stage, (4) the outputs that are obtained at the end of the stage, and (5) a BPMN diagram that graphically shows the activities of the stage [11]. The methodology is resumed in Table 1.

Although when specifying the methodology, we explained in detail each of its stages, some activities can be difficult to perform without a guide that helps the researcher determine how the stages should be developed. On the one hand, the methodology has several stages. While all stages are important for the development of a new set of heuristics, the researcher may be confused regarding how to apply them and what activities to carry out in each one (how to classify the information collected, what information to select, how to document the activities carried out, etc.). On the other hand, the methodology has a formalism that can be difficult to understand for the first time. The use of inputs,

Table 1
Definition, inputs and outputs for each stage of the methodology [11].

Name	Definition	Input	Output
Step 1: Exploratory stage	Perform a literature review	A specific application domain that needs a new set of heuristics or checklist	① Information regarding the application (definitions and features) ② Usability and UX attributes ③ Sets of heuristics and/or other relevant elements
Step 2: Experimental stage	Analyze data that are obtained in different experiments to collect additional information that has not been identified in the previous stage	A specific application domain that needs a new set of heuristics or checklist	④ Additional specific features of the application ⑤ Detected usability problems ⑥ Problems with existing heuristics
Step 3: Descriptive stage	Select and prioritize the most important topics of all information that was collected in the previous stages	① Information regarding the application (definitions and features) ② Usability and UX attributes ③ Sets of heuristics and/or other relevant elements ④ Additional specific features of the application ⑤ Detected usability problems	⑦ Selected information regarding the application ⑧ Selected features of the specific application domain ⑨ Selected usability/UX attributes ⑩ Selected sets of heuristics and/or other relevant elements
Step 4: Correlational stage	Match the features of the specific application domain with the usability/ UX attributes and existing heuristics (and/or other relevant elements)	⑦ Selected information regarding the application ⑧ Selected features of the specific application domain ⑨ Selected usability/UX attributes ⑩ Selected sets of heuristics and/or other relevant elements	⑪ Matched features, attributes and existing heuristics (and/or other relevant elements) ⑫ Categories
Step 5: Selection stage	Keep, adapt and/or discard the existing sets of usability/UX heuristics that were selected in Step 3 (and/or other relevant elements)	⑩ Selected sets of heuristics and/or other relevant elements ⑪ Matched features, attributes and existing heuristics (and/or other relevant elements)	⑬ Classified heuristics (to keep, adapt, create and eliminate)
Step 6: Specification stage	Formally specify the new set of usability/UX heuristics	⑥ Problems with existing heuristics ⑪ Matched features, attributes and existing heuristics (and/or other relevant elements) ⑫ Categories ⑬ Classified heuristics (to keep, adapt, create and eliminate)	⑭ Set of proposed heuristics
Step 7: Validation stage	Validate the set of heuristics through several experiments in terms of their effectiveness and efficiency in evaluating the specific application	⑭ Set of proposed heuristics	⑮ Heuristic evaluation results: effectiveness of heuristics ⑯ Expert judgment results (survey) ⑰ User tests results: users' perceptions
Step 8: Refinement stage	Refine and improve the new set of heuristics based on the feedback that was obtained in Step 7	⑮ Heuristic evaluation results: effectiveness of heuristics ⑯ Expert judgment results (survey) ⑰ User tests results: users' perceptions	⑱ Refining document: (1) What heuristics to create, refine and/or delete, why, and how to do it; (2) What steps to repeat

outputs, tables and BPMN diagrams are a great help in facilitating the heuristic development process, but at the beginning, it can be complex for the researcher to learn how to use them properly. For instance, in some stages, certain inputs and/or outputs are optional depending on the activities that are carried out. These options could confuse the researcher on how to start or finish the stages.

In addition, it is possible to repeat some stages of the methodology to refine and improve the set of proposed heuristics. For the researcher, it may not be evident which stages to repeat, at which time, or when to stop iterating. Due to the above, we believe that it is necessary to present a detailed explanation of how to apply each stage of the methodology. We aim to help the researcher in the heuristic development process using the methodology.

4. How to use the methodology in practice: a user guide

In this section, we explain in detail how to apply each methodology stage and how to iterate to create a new set of usability/UX heuristics for a specific application domain. The researcher can adapt the activities of the stages that he/she considers appropriate. We hope that this detailed explanation will serve as a guide for the researcher to apply the methodology to develop a new set of heuristics. This section is divided into two parts:

1 In Section 4.1, we explain how to apply each stage. For each stage, we show the inputs, activities performed, outputs, and how to document the process. In addition, justifications are presented for each decision made regarding which activities to perform, which inputs/outputs to use, which information to select, etc.
2 In Section 4.2, we explain how to iterate, when to iterate and how many iterations should be carried out to refine and improve the new set of heuristics. We include diagrams to show how to document the iterations.

To explain the process of developing heuristics, the development of a set of UX heuristics for national park websites is used as a case study [12]. This case study is used to explain how to apply the methodology to develop heuristics that evaluate aspects other than usability, such as the UX. In addition, several other case studies are analyzed to complement the explanation. We include examples of other sets of heuristics for e-commerce, grid computing, and smartphones, among others.

4.1. How to use methodology's stages

The explanation of how to apply each stage is presented below.

4.1.1. Step 1: exploratory stage

• ***Input: specific application domain that needs a new set of heuristics or checklist***. First, the researcher identifies a domain, for instance national park websites, smartphones, and E-commerce, that needs a new set of heuristics (or checklist) since there are no sets that effectively evaluate the usability/UX and their attributes. Then, the researcher must gather useful information for the development of heuristics, obtaining three outputs.

• ***Output: ① information regarding the application (definitions and features)***. Different information sources can be reviewed to collect information related to the specific domain. For instance, for national park websites, the researchers collected information regarding (1) the definitions of national parks and national park websites; (2) the purpose of a national park website; (3) the context of use of a national park website; (4) the advantages and disadvantages of using a national park website; and (5) the general and specific features of a national park website. General features are crosswise features for any kind of software product (and for that reason, they are also related to the national park websites). Specific features are features of the specific domain that differentiate them from other types of software products.

The researcher can gather other types of information that he/she considers useful for the development of heuristics, such as:

1 Types or classifications: A domain can be very broad and therefore have different classifications. For instance, for grid computing applications [22], the authors identified four types of grids: computer, data, service, and equipment grids.
2 Applications: Domain application areas. For instance, for grid computing applications [22], the authors identified four application areas: job submission, monitoring, visualization, and web portals.
3 Domain-specific tasks: Tasks that the user performs using products, systems or services of the specific domain. For instance, in [22], the authors have identified seven specific tasks that users perform using grid computing.
4 Target audiences: Users and/or types of users to whom the product, system or service of the specific domain is oriented.

• ***Output: ② usability and UX attributes***. The researcher collects several attributes that define the usability/UX. We recommend reviewing different proposals stated by different authors to cover all the usability/UX dimensions, and then (in step 3) decide which of them is the most appropriate to evaluate the specific domain. For instance, to develop heuristics for national park websites, the researchers collected three proposals (two proposals for the usability and one for the UX): (1) usability attributes defined by the ISO 9241 standard [14]; (2) usability attributes proposed by Nielsen [23]; and (3) UX factors proposed by Morville [16]. The researcher can gather other proposals regarding usability/UX attributes, such as (1) factors that influence the UX [17]; (2) elements that influence the UX [18]; (3) elements of UX [19]; and (4) UX attributes [24].

• ***Output: ③ sets of heuristics and/or other relevant elements***. The researcher collects an existing set of heuristics and/or other useful information (such as principles, design recommendations, guidelines, and patterns, related to the specific domain). The latter can be particularly useful when there are no other sets of related heuristics in addition to Nielsen's heuristics [3].

For example, to develop heuristics for national park websites, the researchers collected two existing sets of heuristics: (1) Nielsen's 10 heuristics (traditional heuristics) [3] and (2) heuristics for virtual museums (specific heuristics) [25]. The researchers did not gather other relevant elements because they considered that with the information collected and the sets of heuristics reviewed, their collection was sufficient. In [26], for the u-learning domain, the authors selected two existing sets of heuristics to develop the new set: (1) Nielsen's 10 heuristics [3] and (2) heuristics for u-learning [27].

If the researcher is developing a set of heuristics for a specific application domain that is recent and/or emerging and for which there are no sets of specific heuristics that can serve as a basis, then we recommend reviewing other information that can provide important elements for creating the new heuristics. For example, in [28], the authors proposed a set of cultural-oriented usability heuristics. Since there were no sets of specific heuristics related to cultural aspects, the authors collected information on cultural dimensions and used those dimensions to define the new set of heuristics.

4.1.2. Step 2: experimental stage

• ***Input: specific application domain that needs a new set of heuristics or checklist***. In this stage, the researcher can analyze data obtained in previous studies or perform their own experiments to collect additional information to develop the new set of heuristics. This step is optional. For instance, to develop heuristics for national park websites, the researchers did not perform this stage because (1) they performed an exhaustive literature review in the previous stage; (2) they found no previous studies related to national park websites' usability/UX; and (3) they did not have resources to perform the experiments.

The researcher decides if he/she needs to perform experiments.

Table 2
Information regarding national park websites prioritized.

Value	Information type	Description
3	Definition of the national park (physical place)	Definition proposed by the National Forestry Corporation (CONAF, Chile)
3	Definition of the national park (physical place)	Definition proposed by U.S. government
2	Definition of the national park (physical place)	Definition proposed by CITMA (website)
2	Purpose and definition of the national park (physical place)	Definition proposed by Real Academia Española [34]

Based on the previous data, the experiments performed (e.g., heuristic evaluation) and the way of interpreting the results, the researcher can obtain three different outputs. Since this stage is optional, the outputs are also optional. Depending on the activities performed, the researcher may have one to three outputs.

• ***Output: ④ additional specific features of the application***. The researcher can analyze data from previous studies or perform new experiments to collect new features not identified in the previous stage. Although this output is optional, it is useful for Step 3: the Descriptive stage. For instance, to develop heuristics for grid computing applications [22], authors analyzed the results obtained by other authors regarding a specific application of grid computing [29] to identify additional specific features. In [28], for e-commerce websites, authors conducted interviews with expert evaluators to identify which elements of e-commerce should be evaluated with a set of heuristics.

• ***Output: ⑤ detected usability problems***. The researcher can analyze data from previous studies or perform new experiments to identify usability problems related to the domain and create heuristics that evaluate and cover those problems. Although this output is optional, it is useful for Step 3: Descriptive stage.

For instance, to develop heuristics for smartphones [30] and cultural aspects [28], the authors performed a guided inspection to identify usability issues based on Nielsen's heuristics [3]. In [30], the authors carefully analyzed the issues that could not be associated with Nielsen's heuristics to identify which heuristics were necessary to create. In [31], for driving simulator applications, the researchers performed a heuristic evaluation to evaluate “SimuDrive” software using Nielsen's heuristics to identify usability issues and then determine which problems are not possible to associate with Nielsen's heuristics.

• ***Output: ⑥ problems with existing heuristics***. The researcher can analyze data from previous studies or perform new experiments to identify which problems occur while understanding existing heuristics and to avoid them. In addition, the researcher can find interesting elements to include in the new set of heuristics by reviewing existing sets. Although this output is optional, it is useful for Step 6: Specification stage.

For instance, to develop heuristics for transactional websites [32,33], the authors analyzed the results obtained in 6 heuristic evaluations performed by other evaluators to determine the problems using Nielsen's heuristics when evaluating the website www.hotelclub.com. This analysis allowed the identification of the present heuristics problems that need to be understood.

4.1.3. Step 3: descriptive stage

• ***Inputs: ① information regarding the application (definitions and features); ② usability and UX attributes; ③ sets of heuristics and/or other relevant elements; ④ additional features of the application; and ⑤ detected usability problems***. For step 3, inputs ①, ②, and ③ are mandatory. Inputs ④ and ⑤ are optional, since they depend on whether the researcher performed step 2 or not. For instance, to develop heuristics for national park websites, the researchers did not work with inputs ④ and ⑤ since step 2 was not performed.

In the descriptive stage, the researcher selects and prioritizes the information collected in steps 1 and 2 (if it applies). The information is selected and grouped into five topics: (1) information regarding the specific domain; (2) features of the specific domain; (3) usability/UX attributes; (4) existing sets of heuristics and/or other relevant elements; and (5) usability problems detected. The last topic (5) is optional, depending on whether the researchers performed step 2. Then, the information is prioritized by each topic using a scale of 3 levels (3: highly important; 2: somewhat important; and 1: not important). As the information is prioritized, the outputs of step 3 are generated.

To develop heuristics for the national park websites, the researchers selected and grouped the information in the corresponding topics (except for topic 5, since they did not perform step 2) [11]. However, they prioritized the information according to what they considered important, without using the scale proposed in the methodology due to time limitations. To explain how to use the scale, we present below how the researchers should have applied the scale for the development of heuristics for national park websites.

• ***Output: ⑦ selected information regarding the application***. The researcher orders and prioritizes the information collected regarding the domain (definitions, classifications, areas of use, etc.) using input ① (information regarding the application).

To develop heuristics for national park websites, the researchers prioritized the information, as shown in Table 2. This table shows the type of information collected, a brief description, and the value assign according to their importance.

Since there was no definition for national park websites, the researchers proposed one. After reviewing and analyzing their work [11], we identified that they proposed the definition using the information qualified with value 3 (see Table 2).

• ***Output: ⑧ selected features of the specific application domain***. The researcher orders and prioritizes the information collected regarding the features of the specific domain using inputs ① (information regarding application) and ④ (additional features of the application, if it applies). The researcher can group the features into general (crosswise features for any kind of software product) and specific features (features of the specific domain that differentiate them from other types of software products).

To develop heuristics for the national park websites, the researchers identified 8 general features and 9 specific features [11]. After reviewing and analyzing their work [12], we detected the following:

1 The researchers selected all general features collected in step 1, which means that they had to rate each feature with value 3 (the highest, indicating that all general features are important to consider for the development of heuristics).
2 The researchers selected 8 of 9 specific features collected in step 1, which means that they had to rate 8 specific features with value 3 (the highest) and 1 specific feature with value 1 (the lowest, indicating that this feature is not important, e.g., because it is not related to the domain).

• ***Output: ⑨ selected usability/UX attributes***. The researcher orders and prioritizes the information collected regarding usability/UX attributes using input ② (usability and UX attributes). The researcher can prioritize the usability and UX attributes separately or together. Additionally, the researcher can prioritize the attributes by following two approaches (or the mixture of both):

1 Make a list of all the identified attributes, regardless of the author

Table 3
Prioritizing sets of attributes.

	Value	Author of the proposal	Amount of attributes	Proposal of attributes	Justification
Usability	3	Nielsen [23]	5	Learnability, efficiency, memorability, errors and satisfaction.	Since Nielsen proposal includes more usability attributes than the ISO standard, a value of priority 3 and 2 has been assigned.
	2	ISO 9214 [14]	3	Effectiveness, efficiency and satisfaction.	
UX	3	Morville [16]	7	Useful, usable, desirable, findable, credible, accessible and valuable.	Only the UX factors proposed by Morville were reviewed.

that proposed them.

2 Make a list of a set of attributes, proposed by different authors.

That is, the researcher decides if he/she wants to prioritize considering all the attributes or if he/she wants to prioritize the proposals of attributes suggested by a particular author. Regardless of the approach chosen to prioritize, the researcher must justify why he/she has chosen certain attributes and why he/she has discarded others.

For instance, to develop heuristics for national park websites, the researchers collected three proposals of attributes for the usability/UX (two proposals for usability and one for UX, see step 1). After reviewing and analyzing their work [12], we detected that the researchers performed a prioritization at two levels. First, they prioritized the set of attributes proposed by different authors for usability and UX (see Table 3), and then they selected which attributes of each proposal will be used for the development of the heuristics (see Table 4). Tables 3 and 4 show how the researchers prioritized the information using the scale proposed in the methodology (column named "value"). In addition, these tables also show the justification of why certain proposal/attributes have been selected and/or discarded (column named "justification").

• ***Output:*** ⑩ ***selected sets of heuristics and/or other relevant elements***. The researcher orders and prioritizes the information collected regarding the sets of heuristics and/or other relevant elements using inputs ③ (sets of heuristics and/or other relevant elements) and ⑤ (detected usability problems, if it applies).

The researcher can select one or more sets of existing heuristics. On the one hand, if there are no specific sets to evaluate similar domains, then we recommend using Nielsen's heuristics [3] as a basis. On the other hand, if there are other similar sets, then the researcher can choose those that he/she considers useful (the scale to prioritize can help to define which set is the most appropriate). In addition, if there are no similar sets of heuristics, then other relevant elements collected can serve as a basis to develop the new heuristics (patterns, guidelines, etc.). For example, usability problems detected in step 2 can help to identify which heuristics need to be created to detect usability problems related to the specific domain. The researcher makes a list of the detected usability problems and determines which elements of the domain should be evaluated with the new set of heuristics.

For instance, to develop heuristics for national park websites, the researchers identified two sets of heuristics: Nielsen's heuristics (traditional heuristics) [3] and heuristics for virtual museums (specific heuristics) [25]. No other relevant elements were collected. The researchers selected heuristics for virtual museums since Nielsen's heuristics are too generic and do not cover some specific features of national park websites. This means that the researchers had to rate heuristics for virtual museums with value 3 (the highest) and Nielsen's heuristics with value 1 (the lowest).

In [28], to develop cultural-oriented usability heuristics, the authors identified three sets of heuristics (1 set of traditional heuristics and 2 sets of specific heuristics) and other relevant elements (cultural dimensions and cultural aspects in websites, models and guidelines). Although authors in [28] did not prioritize the information collected using the scale proposed in our methodology, they prioritized the importance of the information. Table 5 shows what the values assigned to each element would be based on the selection made by the authors in [28] and its justification.

Depending on the specific domain, a researcher can find several sets of existing heuristics that may be used for the development of new heuristics. If the researcher finds a large number of sets of existing heuristics, then it can be difficult for him/her to prioritize the sets. To facilitate the prioritization, we recommend performing a comparative analysis between sets of heuristics first and then prioritizing them. The researcher can review the heuristics of each set and match those that evaluate the same aspects. Table 6 shows an example of how to analyze different sets of heuristics for the e-commerce website domain. As

Table 4
Prioritizing attributes individually.

	Value	Attribute	Author of the proposal	Justification
Usability	3	Learnability	Nielsen [23]	The 5 attributes proposed by Nielsen were selected since they are more complete than those proposed by the ISO standard.
	3	Efficiency		
	3	Memorability		
	3	Errors		
	3	Satisfaction		
UX	3	Useful	Morville [16]	Six of the seven factors proposed by Morville were selected. "Accessible" was not considered, as it is not intended to evaluate the accessibility with the new set of heuristics.
	3	Usable		
	3	Desirable		
	3	Findable		
	3	Credible		
	1	Accessible		
	3	Valuable		

shown in Table 6, based on the comparative analysis, it is possible to identify aspects that are evaluated by a single set of heuristics or by several of them. In addition, this analysis allows the determination of whether the existing sets of heuristics evaluate all the features of the specific domain in study. It is even possible that after reviewing the existing sets of heuristics, the researcher determines that a certain set properly evaluates the specific application domain, and therefore, he/she may stop the process of developing new heuristics.

4.1.4. Step 4: correlational stage

• ***Inputs: ⑦ selected information regarding the application; ⑧ selected features of the specific application domain; ⑨ selected usability/UX attributes; and ⑩ selected sets of the heuristics and/or other relevant elements***. In this stage, the researcher performs two activities: (1) Match the features of the specific application domain, the usability/UX attributes, and the existing heuristics (and/or other relevant elements); and (2) Group heuristics into categories (if it applies).

• ***Output: ⑪ matched features, attributes and existing heuristics (and/or other relevant elements)***. The researcher performs the feature match using all the inputs (⑦, ⑧, ⑨, and ⑩). To perform the match, the researcher first creates a list with all the features of the specific domain. Then, he/she matches each feature with the usability/UX attribute. Finally, the researcher analyzes whether the existing heuristics selected completely or partially cover each of the features and usability/UX attributes. It is normal that some features are not covered by an existing heuristic (this means that it will be necessary to create a heuristic to evaluate those features).

For instance, to develop heuristics for national park websites, the researchers matched the general and specific features of national park websites, the usability/UX attributes, and the set of heuristics for virtual museums [11]. Researchers considered both the usability and UX attributes, and for this reason, each feature has associated usability and/or UX attributes. In addition, some existing heuristics partially cover some features and attributes, while other features and attributes are not covered by any existing heuristic.

In [28], to develop cultural-oriented usability heuristics, the authors performed a match among cultural dimensions (other relevant element), features and the heuristics created. Authors did not match the usability attributes, but they did use the usability attributes proposed by Nielsen to create the new heuristics [23]. Based on the information reviewed in [28], it was possible to determine how the authors produced the matching. Table 7 shows the match among features, usability attributes, Nielsen's heuristics and cultural dimensions. As shown in Table 7, it is possible to make a match with other relevant elements (in this case, cultural dimensions) in addition to the existing heuristics.

• ***Output: ⑫ categories***. The researcher decides whether the new heuristics will be grouped into categories or will not be based on the features identified (it is not mandatory to define categories).

Depending on the specific application domain, it may be useful to group heuristics into categories. Each category can group heuristics that evaluate certain specific aspects. To create and/or name the categories, the output ⑪ obtained in this same stage can be useful.

For instance, to develop heuristics for national park websites, the researchers did not define categories since each heuristic evaluates a different feature of the specific domain. In [22], to develop heuristics for grid computing applications, authors classified the heuristics into three categories: (1) Design and aesthetics heuristics; (2) Flexibility and navigation heuristics; and (3) Errors and help heuristics. To develop playability heuristics for mobile games [41], the authors grouped the heuristics into three categories: (1) Game usability heuristics; (2) Mobility heuristics; and (3) Gameplay heuristics. In [42], the authors created four categories to group the heuristics for educational games: (1) Interface heuristics; (2) Content heuristics; (3) Pedagogical/educational heuristics; and (4) Multimedia heuristics.

4.1.5. Step 5: selection stage

• ***Inputs: ⑩ selected sets of heuristics and/or other relevant elements; and ⑪ matched features, attributes and existing heuristics (and/or other relevant elements)***. In the selection stage, the researcher reviews existing heuristics and determines which heuristics to keep, adapt, create and/or eliminate to develop the new set of heuristics. For the above, the researcher uses input ⑩ to review each existing heuristic and input ⑪ to identify which new heuristics are necessary to create.

• ***Output: ⑬ classified heuristics (to keep, adapt, create and eliminate)***. The researcher classifies the heuristics using all the inputs. Using input ⑩ (selected sets of heuristics and/or other relevant elements), the researcher reviews each existing heuristic and decides which heuristics to keep, adapt, or eliminate. When reviewing different sets of heuristics, it is possible to identify very similar heuristics obtained from different sets. In this case, it is necessary to eliminate the redundancy and complement the heuristic's information if necessary but without repeating it.

Using input ⑪ (matched features, attributes and existing heuristics, and/or other relevant elements), the researcher can identify which new heuristics are necessary to create to evaluate specific domain-related features of the specific domain. If the researcher analyzed other relevant elements (already included in output ⑪), then these can be helpful in determining which new heuristics to create.

For instance, the selection process performed to develop the heuristics for national park websites is synthesized in [11] and presented in detail in [12]. The authors analyzed the existing set of 15 heuristics for virtual museums but did not analyze other relevant elements. Using

Table 5
Sets of heuristics and other relevant elements prioritized for cultural-oriented usability heuristics [28].

Type of information	Description	Details	Value	Justification
Sets of heuristics	Nielsen's heuristics [3]	10 usability heuristics	3	10 Nielsen's heuristics were selected since authors did not find specific cultural-oriented usability heuristics
	Quiñones et al. heuristics [32]	13 usability heuristics for transactional web applications	1	This set of heuristics was not selected since they do not evaluate cultural aspects
	Bonastre and Granollers heuristics [35]	64 usability heuristics for e-commerce websites (presented as questions)	1	This set of heuristics was not selected since the study did not present experimental validation
Other relevant elements	Hofstede's cultural dimensions [36]	1. Power distance dimension (PD)	3	PD, IDV, MAS, and UA dimensions were selected since the authors considered that they help to differentiate cultures [28]
		2. Individualism dimension (IDV)	3	
		3. Masculinity dimension (MAS)	3	
		4. Uncertainty avoidance dimension (UA)	3	
		5. Long-term orientation dimension	1	Dimensions 5 and 6 were not selected since they do not add much practical information and applicability to the different cultures [28]
		6. Indulgence and restraint dimension	1	
	Cultural aspects in websites	Model for developing usable cross-cultural websites (Smith et al. [37])	3	The 3 proposals were selected since all of them provide relevant information regarding cultural aspects for different types of websites [28]
		Guidelines for cultural-oriented interfaces (cross-cultural effects in e-retailing, Huggins et al. [38])	3	
		Guidelines for cultural-oriented interfaces (effects of usability and web design attributes on user preference for an e-commerce website, Sangwon et al. [39])	3	

input ⑩, the researchers reviewed 15 heuristics, of which 0 heuristics were kept unchanged, 12 were adapted, and 3 heuristics were eliminated, as they were not applicable to the evaluation of national park websites. Using output ⑪, the authors determined that it was necessary to create 6 new heuristics to evaluate the specific features of the national park websites.

In [28], to develop cultural-oriented usability heuristics, the authors also performed a heuristic selection process, although they did not document it, as it is proposed in our methodology. Based on the information reviewed in [28], it was possible to determine how the authors would have performed the selection process. Table 8 shows some examples of the heuristic selection process for creating the set of heuristics using the existing set of Nielsen's heuristics [3]. The authors reviewed 10 heuristics, and all were adapted (none were kept or eliminated). Each heuristic was adapted to evaluate related cultural aspects, and its specification was detailed. In addition, the investigators used cultural dimensions to develop a new set of heuristics (as “other relevant elements”). Based on the above, these authors created 2 new heuristics to evaluate specific features.

4.1.6. Step 6: specification stage

• ***Inputs:* ⑥ *problems with existing heuristics;* ⑪ *matched features, attributes and existing heuristics (and/or other relevant elements);* ⑫ *Categories; and* ⑬ *classified heuristics (to keep, adapt, create and eliminate)***. In the specification stage, the researcher defines the amount of heuristics; determines the elements to include in the heuristics' specification; groups the heuristics into categories (if it applies); and formally specifies the new set of heuristics using a template. We recommend using a standard template because all the heuristics will be consistent in the information they present, and each heuristic will have a level of detail that helps to understand it better. The researcher may decide whether to use the complete template or just some elements to specify the heuristics.

We recommend keeping the number of heuristics relatively small (between 10 and 16) [20], and if it is necessary to add a greater level of detail, then we suggest developing an additional checklist. The standard template proposed to specify the heuristics can be reviewed in [11].

• ***Output:* ⑭ *set of proposed heuristics***. To specify a new set of heuristics, inputs ⑪ and ⑬ are mandatory. Inputs ⑥ and ⑫ are optional since input ⑥ depends on whether the researcher performed step 2 or not, and input ⑫ depends on whether the researcher creates categories in step 4.

For instance, to develop heuristics for national park websites, the researchers did not work with inputs ⑥ and ⑫ because step 2 was not performed and no categories were defined to group the heuristics in step 4. Using inputs ⑪ and ⑬, the researchers created 14 heuristics and used the 12 elements of the template to specify the heuristics. The set of heuristics created can be reviewed in [11] (the complete version can be reviewed in [12]).

In [30], to develop heuristics for smartphones, the authors develop 12 heuristics; and in [22], the authors develop 10 heuristics for grid computing applications, grouping the heuristics into three categories. In both articles, the authors specify heuristics using the following 7 elements: ID, name, definition, explanation, examples, benefits and problems associated with misinterpretation. In both cases, the authors used the methodology proposed by Rusu et al. [43] to create the new set of heuristics. Our methodology is based on [43], so the template is similar. However, we have added more elements to the specification template.

4.1.7. Step 7: validation stage

- ***Inputs:* ⑭ *Set of proposed heuristics***

In this stage, the researcher validates the new set of proposed heuristics through several experiments. These experiments may check

Table 6
Comparative analysis between four sets of heuristics to evaluate transactional websites.

Nielsen's heuristics [3]	Paz et al. heuristics [40]	Quiñones et al. heuristics [33]	Díaz et al. heuristics [28]
N1: Visibility of the system status	T1: Visibility and clarity of the system's elements T2: Visibility of the system status	SWT1: Visibility of the system status	CH1: Visibility of the system status
N2: Match between the system and the real world	T3: Match between the system and user's cultural aspects	SWT5: Match between the system and the real world	CH2: Match between the system and the real world
–	T4: Feedback of the transaction	SWT2: Keep the user informed regarding the status of transactions	–
N4: Consistency and standards	T5: Alignment to the web standards design T7: Standard iconography	SWT8: Use of standards and web symbology	CH4: Consistency and standards
	T6: Consistency of design	SWT7: Consistency between the elements of the system	
–	–	SWT4: Security and speed of the transactions	–
–	–	–	CH11: The information structure

the effectiveness and efficiency of the new set of heuristics when evaluating the usability/UX of specific domains. For each experiment, input ⑭ is used. Based on the experiment performed, three outputs can be obtained:

1 Experiment: Heuristic evaluations. Output: ⑮ Heuristic evaluation results: effectiveness of heuristics.
2 Experiment: Expert judgments. Output: ⑯ Expert judgment results (survey): utility, clarity, ease of use, need for checklist and comments regarding each heuristic.
3 Experiment: User tests. Output: ⑰ User tests results: users' perceptions.

For instance, to develop heuristics for national park websites, the researchers validated the heuristics through heuristic evaluation and expert judgment (survey) in the first iteration and through user tests (co-discovery and focus group) in the second iteration.

In [26], to establish heuristics for u-learning applications, the authors validated the heuristics in two iterations through expert judgment (survey) and two heuristic evaluations. In [31], to create heuristics for driving simulators, authors validated the heuristics through expert opinion and three heuristic evaluations.

To develop cultural-oriented heuristics [28], the authors validated the heuristics through two heuristic evaluations in the second iteration, through three heuristic evaluations in the third iteration, and through expert judgment (survey) in the fourth iteration.

4.1.7.1. Validation through heuristic evaluation. The researcher checks the new set of heuristics against control heuristics (traditional or specialized heuristics). For this, the researcher (1) selects applications to evaluate; (2) selects the set of control heuristics; (3) selects evaluators for a heuristic evaluation with similar experience (define control and experimental groups); and (4) evaluates the effectiveness of the heuristics in terms of the 5 criteria. The new set of usability/UX heuristics performs well, and it is an effective instrument when better results than the control heuristics are obtained in terms of the 5 criteria [11]. Table 9 shows how different authors validate the set of proposed heuristics through heuristic evaluations.

The evaluations involved between 2 and 5 evaluators per group. Some evaluations involved more than one control or experimental group, so column "Number of evaluators involved" shows more evaluators in certain cases. In addition, in some evaluations, the authors worked with only the experimental group (without the control group) because it was the last iteration (heuristics for smartphones [30], cultural-oriented heuristics [28]). The last column of the table shows six criteria used by different authors to validate the heuristics. Each criterion is detailed below.

1 C1: Numbers of correct and incorrect associations of problems to heuristics
2 C2: Number of usability/UX problems identified
3 C3: Number of specific usability/UX problems identified
4 C4: Number of identified usability/UX problems that qualify as more severe
5 C5: Number of identified usability/UX problems that qualify as more critical
6 C6: Average of Hofstede' cultural dimensions.

Table 7
Match among features, usability attributes, Nielsen's heuristics and cultural dimensions [28].

Feature	Usability attribute	Nielsen's heuristic [3]	Cultural dimension
Informed actions – status changes – feedback	Memorability	N1: Visibility of the system status	(UA) Uncertainty avoidance
Familiar concepts – conventions	Memorability	N2: Match between the system and the real world	(IDV vs COL) Individualism vs. collectivism
Control – undo/redo	Efficiency	N3: User control and freedom	(MAS vs FEM) Masculinity vs. femininity
Conventions – standards	Satisfaction	N4: Consistency and standards	(IDV vs COL) Individualism vs. collectivism
Avoid mistakes – detailed design – prevention	Errors	N5: Error prevention	(UA) Uncertainty avoidance
Memory load – remember information – instructions	Memorability	N6: Recognition rather than recall	(UA) Uncertainty avoidance
Efficient – simple – efficacy	Efficiency	N7: Flexibility and efficiency of use	(MAS vs FEM) Masculinity vs. femininity
Clean interface – significant elements	Satisfaction	N8: Aesthetic and minimalist design	(IDV vs. COL) Individualism vs. collectivism
Clear content – suggesting solution – details	Errors	N9: Help users recognize, diagnose, and recover from errors	(UA) Uncertainty avoidance
Documentation – information – adequate	Learnability	N10: Help and documentation	(UA) Uncertainty avoidance
Hierarchic – organization – options	Efficiency	–	(PD) Power distance
Goal oriented – concise	Learnability	–	(MAS vs FEM) Masculinity vs. femininity

Table 8
Heuristic selection process for creating cultural-oriented heuristics.

ID	Heuristic Name/Explanation	Action	Set of existing heuristics	Aspect or feature covered	Applicability
N3	User control and freedom	Adapt	Nielsen's heuristics [3]	Control – undo/redo (partially covered)	(3) Critical
N7	Flexibility and efficiency of use	Adapt	Nielsen's heuristics [3]	Efficient – simple –efficacy (partially covered)	(3) Critical
CH11	The information structure	Create	–	Hierarchic – organization – options	(3) Critical
CH12	Accurate and detailed results	Create	–	Goal oriented – concise	(3) Critical

Table 9
How different authors validate a new set of heuristics through heuristic evaluations.

Set of proposed heuristics	Control heuristics selected	Number of evaluators involved		Criteria used to evaluate					
		Control group	Experimental group	C1	C2	C3	C4	C5	C6
UX heuristics for national park websites [11]	Usability heuristics for virtual museums [25]	3	3	x	x	x	x	x	
Usability heuristics for u-learning applications [26]	Nielsen's heuristics [3]	2	2		x		x	x	
Usability heuristics for driving simulators [31]	Nielsen's heuristics [3]	3	3		x		x	x	
Cultural-oriented usability heuristics [28]	Nielsen's heuristics [3]	3	3		x	x	x	x	x
	–	0	21–27		x	x	x	x	x
Smartphone's usability heuristics [30]	Nielsen's heuristics [3]	2	2		x	x	x		
	Nielsen's heuristics [3]	3	3		x	x	x		
	–	0	27		x	x	x		
Usability heuristics for grid computing applications [22]	Nielsen's heuristics [3]	4	4		x	x	x		
	Nielsen's heuristics [3]	3	3		x	x	x		
Usability heuristics for transactional websites [33]	Nielsen's heuristics [3]	19	4		x	x	x		

The first 5 criteria have been proposed in our methodology [11]. The last criterion (C6) is specific for the area of application (cultural aspects [28]), so it only applies to that case study.

- ***Output:⑮Heuristic evaluation results: effectiveness of heuristics***

The researcher evaluates the effectiveness of the new set of heuristics based on 5 criteria. Each criterion is explained below.

1 Numbers of correct and incorrect associations of problems to heuristics

The new set of heuristics is an effective instrument when it has fewer incorrect associations of the identified usability/UX problems to the heuristics than that of the control heuristics.

The researcher must perform a critical and constructive analysis of the results that are obtained by the control group and experimental group. The researcher must (critically, without influencing the results) analyze the associations of the problems that are detected with the heuristics and determine whether each association is correct or incorrect.

This analysis provides qualitative data, in which it is possible to identify why the evaluators incorrectly associated the problems with certain heuristics and solved the problems (if any). The following problems are considered:

1 The definition of a heuristic is unclear. It is necessary to improve its definition.
2 There are heuristics with similar definitions. It is necessary to clearly differentiate the heuristics and the aspects that they evaluate.
3 There are heuristics that evaluate the same aspect (redundancy). It is necessary to eliminate a heuristic.
4 The evaluators did not understand how to use the heuristics. The heuristics are well specified, but the evaluators did not know how to use them correctly.
5 Some heuristics do not have associated problems. These heuristics should not be discarded but should be analyzed with respect to why they do not have associated problems.

To measure the effectiveness of the new set of heuristics compared to the control heuristics, four percentages must be calculated using the following Eq. (1):

$$CA = \frac{\sum_{n=1}^{T} CAHn}{TP} \times 100 \quad (1)$$

where

- CA: correct associations;
- T: total number of heuristics of the set;
- CAHn: number of correct associations of the problems to the heuristic “n”;
- TP: total usability/UX problems identified.

Two percentages should be calculated:

1 Percentage of correct associations of problems to the new set of heuristics (CA1);
2 Percentage of correct associations of problems to the control heuristics (CA2).

The new heuristics perform well when CA1 is greater than or equal to CA2, that is, when the percentage of correct associations of the new set of heuristics is greater than or equal to the percentage of correct associations of the control heuristics.

The effectiveness in terms of the number of incorrect associations of the problems to the heuristics can be represented as a percentage by the following Eq. (2):

$$IA = \frac{\sum_{n=1}^{T} IAHn}{TP} \times 100 \quad (2)$$

where:

- IA: incorrect associations;
- *T*: total number of heuristics of the set;
- IAHn: number of incorrect associations of the problems to the heuristic “*n*”;

- TP: total usability/UX problems identified.

Two percentages should be calculated:

1 Percentage of incorrect associations of the problems to the new set of heuristics (IA1);
2 Percentage of incorrect associations of the problems to the control heuristics (IA2).

The new heuristics perform well when IA1 is less than IA2, that is, when the percentage of incorrect associations of the new set of heuristics is less than the percentage of incorrect associations of the control heuristics. Formulas (1) and (2) are applicable when each problem is associated with only one heuristic. When a problem is associated with more than one heuristic, CA and IA are difficult to interpret.

2 Number of usability/UX problems identified

Three categories of problems are expected:

1 (P1) Problems that are identified by both groups of evaluators (common problems identified by both groups),
2 (P2) Problems that are identified only by the group that used the new set of heuristics (without considering the common problems),
3 (P3) Problems that are identified only by the group that used control heuristics (without considering the common problems).

The new heuristics perform well when (P1) and/or (P2) include the highest percentage of problems. A question arises when (P3) has the highest percentage of problems: Why are these problems not identified when using the new set of heuristics? There are two possible reasons:

1 The new heuristics are not able to identify these problems, either because there are no appropriate heuristics or because the heuristics are not properly specified.
2 Evaluators who used the new heuristics subjectively ignored the problems.

Suppositions 1 and 2 may be validated or rejected by expert judgment, complementary evaluations and/or user tests:

1 Through expert judgment, it is possible to determine whether heuristics are poorly defined and/or misunderstood or whether it is necessary to include a new heuristic to evaluate certain aspects of the application.
2 By performing a user test, it is possible to determine whether usability problems that are detected only by the group that used the control heuristics are or are not perceived as real problems by users.

To avoid the scenario in which evaluators subjectively ignore usability problems, the following steps are suggested:

1 Check if the new set of heuristics covers all the attributes of usability/UX and application features that are defined in Step 4: Correlational stage.
2 Instruct (or remind) the evaluators to analyze together the detected usability/UX problems. Evaluators should review each usability/UX problem and determine if all aspects of usability/UX and application features have been evaluated. (It is not possible to completely eliminate subjectivity from the experiment, which is related to the evaluators' experience and the correct use of heuristics).

Explanations through case studies on how to apply this criterion can be reviewed in [22,28,30].

3 Number of specific usability/UX problems identified

The new set of heuristics is an effective instrument if more specific usability/UX problems – related to the evaluated application – are identified compared to the control heuristics (i.e., usability/UX issues that are related to specific features or specific aspects of the application).

The effectiveness in terms of the number of specific usability/UX problems that are identified (ESS) can be represented as a percentage by using the following Eq. (3):

$$\mathrm{ESS} = \frac{\mathrm{NSP}}{\mathrm{TP}} \times 100 \tag{3}$$

where:

- ESS: effectiveness;
- NSP: number of specific usability/UX problems identified;
- TP: total usability/UX problems identified.

Two effectiveness values should be calculated:

1 The effectiveness using problems that were identified only by the group that used the new set of heuristics (ESS1).
2 The effectiveness using problems that were identified only by the group that used control heuristics (ESS2).

The new heuristics perform well when ESS1 is at least equal to (but hopefully greater than) ESS2, that is, when the new set of heuristics finds more specific usability/UX problems than the control heuristics.

4 Number of identified usability/UX problems that qualify as more severe

In addition, the new set of heuristics is an effective instrument if it finds more usability/UX problems that qualify as the most severe. In a heuristic evaluation, after preparing the unique and complete list of usability/UX problems, each evaluator separately qualifies the

Table 10
Rating scales of detected usability/UX problems.

Rating scale	Description	Values	Scale
Frequency	Frequency indicates how often the detected usability/UX problem occurs	0–4	**4:** >90% **3:** 51–90% **2:** 11–50% **1:** 1–10% **0:** <1%
Severity [44]	Severity indicates how catastrophic the detected usability/UX problem is	0–4	**4:** Catastrophic problem **3:** Major problem **2:** Minor problem **1:** "Cosmetic" problem **0:** Not a problem
Criticality	Sum of the frequency and severity values.	0–8	–

Table 11
The effectiveness of the new set of heuristics for national park websites.

	Experimental group	Control group	Observations
Set of heuristics used	Heuristics for national park websites (NPH)	Heuristics for virtual museums (VMH)	–
Amount of heuristics (T)	18	15	–
Total of problems identified (TP)	46	47	–
Number of specific problems identified (NSP)	12	9	–
Number of problems identified and qualified with a severity greater than 2 (NPV)	28	21	–
Number of problems identified and qualified with a criticality greater than 4 (NPC)	27	24	–
Problems identified by both groups (P1)	7		Given that (P1) and/or (P2) do not include the highest amount of problems, it is necessary to perform additional experiments to review and improve NPH.
Problems identified by the experimental group (P2)	39	–	
Problems identified by the control group (P3)	–	40	
Total of the correct associations (ΣCAHn)	32	28	–
Total of the incorrect associations (ΣIAHn)	14	19	–
Percentage of the correct associations (CA)	CA1 = 69.6%	CA2 = 59.6%	Given that CA1 > CA2 it is concluded that NPH works better than VMH (NPH has a higher percentage of correct associations)
Percentage of the incorrect associations (IA)	IA1 = 30.4%	IA2 = 40.4%	Given that IA1 < IA2 it is concluded that NPH works better than VMH (NPH has a lower percentage of correct associations)
Effectiveness in terms of number of specific problems identified (ESS)	ESS1 = 26.1%	ESS2 = 19.1%	Given that ESS1 > ESS2 it is concluded that NPH works better than VMH (NPH finds more specific usability/UX problems than VMH)
Effectiveness in terms of number of problems identified and qualified with a severity greater than 2 (ESV)	ESV1 = 60.8%	ESV2 = 44.6%	Given that ESV1 > ESV2 it is concluded that NPH works better than VMH (NPH finds that more usability/UX problems qualify as more severe than VMH)
Effectiveness in terms of number of problems identified and qualified with a criticality greater than 4 (ESC)	ESC1 = 58.6%	ESC2 = 51%	Given that ESC1 > ESC2 it is concluded that NPH works better than VMH (NPH finds that more usability/UX problems qualify as more critical than VMH)

problems. We qualify the problems according to their frequency, severity, and criticality (see Table 10). In this case, the effectiveness depends on the severity since usability/UX problems with high severity impede the proper functioning of the application, thereby generating errors.

The rating scale for severity ranges from 0 to 4, where 0 indicates that the problem is not severe and 4 indicates that the problem is catastrophic. In this sense, it is important to identify those problems that have a severity value that is greater than 2, as this is the average value on the scale of severity. If the researcher decides to use another rating scale for severity, then the average value of that scale should be used in the formula that is presented below. The effectiveness in terms of the number of identified usability problems that qualify as more severe (ESV) can be represented as a percentage by using the following Eq. (4):

$$\mathrm{ESV} = \frac{\mathrm{NPV}}{\mathrm{TP}} \times 100 \tag{4}$$

where:

- ESV: effectiveness;
- NPV: number of usability/UX problems identified qualified with a severity greater than 2;
- TP: total usability/UX problems identified.

Two effectiveness values should be calculated:

1 The effectiveness using problems that were identified only by the group that used the new set of heuristics (ESV1).
2 The effectiveness using problems that were identified only by the group that used control heuristics (ESV2).

The new heuristics perform well when ESV1 is at least equal to (but hopefully greater than) ESV2. That is, the new set of heuristics finds more usability/UX problems that have higher severity scores than the control heuristics.

5 Number of identified usability/UX problems that qualify as more critical

Finally, the new set of heuristics is an effective instrument if it finds more usability/UX problems that qualify as the most critical. If a problem occurs less frequently, then it rarely occurs in the application. This aspect implies that if the problem is also severe, it does not significantly influence the use of the application. When assessing criticality, the frequency and severity of a problem are considered to determine how critical the problem is.

The rating scale for criticality ranges from 0 to 8 (see Table 10). The criticality is obtained by summing the frequency and severity value of the problem. In this sense, it is important to identify those problems that have a criticality value that is greater than 4, as this is the average value on the scale of criticality. If the researcher decides to use another rating scale for criticality, then the average value of that scale should be used in the formula that is presented below. The effectiveness in terms of the number of identified usability/UX problems that qualify as more critical (ESC) can be represented as a percentage by using the following Eq. (5):

$$\mathrm{ESC} = \frac{\mathrm{NPC}}{\mathrm{TP}} \times 100 \tag{5}$$

where:

- ESC: effectiveness;
- NPC: number of usability/UX problems identified qualified with a criticality greater than 4;
- TP: total usability/UX problems identified.

Two effectiveness values should be calculated:

Table 12
How different authors validate a new set of heuristics through expert judgment.

Set of proposed heuristics	Participants selected	Elements evaluated
UX heuristics for national park websites [11]	3 experts performing a heuristic evaluation (practitioners)	− The survey proposed in the methodology was applied [11]. − Heuristics were evaluated concerning 4 dimensions (D1 – Utility, D2 – Clarity, D3 – Ease of use, D4 – Necessity of additional checklist) and 3 questions (Q1 – Easiness, Q2 – Intention, and Q3 – Completeness).
Usability heuristics for u-learning applications [26]	4 expert evaluators	− A semi-structured survey was applied. − Heuristics were evaluated concerning 3 dimensions ((D1 – Utility, D2 – Clarity, and D3 – Ease of use) and 3 open questions (Q1 – Easiness, Q2 – Intention, and Q3 – Completeness).
Cultural-oriented usability heuristics [28]	12 evaluators	− A questionnaire was applied. The model of adoption of methods [45] and a questionnaire proposed in [46] were used to design the survey. − Heuristics were checked concerning 3 dimensions (EUP – Easiness of Use Perceived, UP – Utility Perceived, and UI – and Use Intention).
Smartphone's usability heuristics [30]	Reference [47] does not indicate the number of experts	− An interview was conducted with several experts. − Each heuristic's definition was checked for understandability, clarity and consistency.
	27 evaluators [30] (11 evaluators with no previous experience and 16 evaluators with previous experience)	− A survey was applied. − Heuristics were evaluated concerning 4 dimensions (D1 – Utility, D2 – Clarity, D3 – Ease of use, and D4 – Need of additional evaluation elements – “checklists”).
Usability heuristics for transactional websites [33]	4 experts (practitioners)	− A survey was applied. − Heuristics were evaluated concerning 4 dimensions (D1 – Utility, D2 – Clarity, D3 – Ease of use, and D4 – Relevance).

1 The effectiveness using problems that were identified only by the group that used the new set of heuristics (ESC1).
2 The effectiveness using problems that were identified only by the group that used the control heuristics (ESC2).

The new heuristics perform well when ESC1 is at least equal to (but hopefully greater than) ESC2, that is, when the new set of heuristics finds more usability/UX problems with higher criticality scores than the control heuristics.

Table 11 synthetizes the empirical data that show the effectiveness of the new set of heuristics for national park websites.

4.1.7.2. Validation through expert judgment. The researcher checks the validity of the heuristics by asking experts regarding their appropriateness for evaluating the specific application domain and their perception over the new set of heuristics. For this, the researcher: (1) selects participants for expert judgment (researchers and practitioners) and (2) performs the expert judgment through a survey. The explanation of how to design the survey can be reviewed in [11]. Table 12 shows how different authors validate the set of proposed heuristics through expert judgment.

- ***Output: ⑯ Expert judgment results (survey): utility, clarity, ease of use, need for checklist and comments regarding each heuristic***

The researcher applies the survey and interprets the results obtained. To analyze the results, the researcher can use descriptive statistics for dimensions and questions and/or evaluate the relation between the dimensions and questions. If the researcher decides to apply a different survey to the one we propose in [11], then the analysis should be done accordingly. Explanations regarding how to apply the proposed survey and how to interpret the results can be reviewed in [48–50].

Applying surveys allows the researcher to obtain both quantitative and qualitative results. It is important to include open questions, as the expert can provide comments and relevant qualitative information to refine and improve the set of heuristics (for example, definitions that are difficult to understand, heuristics that should be added, refined or eliminated, and so on).

4.1.7.3. Validation through user tests. The researcher can perform user tests to:

1 Check whether the usability/UX problems identified in a heuristic evaluation using the new set of heuristics are real usability/UX problems for users.
2 Obtain the perceptions of users regarding a specific topic (such as the relevant features of the application domain).

User tests make it possible to determine why the experimental group did not detect the same problems that the control group detected. The tasks to be performed in the user test should involve those problems not detected by the experimental group. Some of the reasons for the above can be:

1 The new heuristics are not properly specified (the experimental group did not find the problems detected by the control group since the heuristics are not well understood, so the heuristics need refining);
2 There are no appropriate heuristics to evaluate some kinds of problems (a heuristic is missing to detect usability problems, so a new heuristic must be created);
3 For the experimental group, the usability problem detected by the control group was not truly a usability problem.

On the other hand, it is possible that during the user test, new usability problems were not detected in the heuristic evaluations. In this case, it is necessary to analyze why these problems were not identified in the heuristic evaluations (poor heuristic definitions, no heuristic to detect the problem, the evaluators subjectively ignored the problems in the evaluation, and so on). As a researcher, it is necessary to review if there is a heuristic in the proposed new set that covers the problem detected in the user test. If the heuristic exists, the researcher should review the heuristic definition and improve it. If the heuristic does not exist, it must be created. Table 13 shows how different authors validate

Table 13
How different authors validate a new set of heuristics via user tests.

Set of proposed heuristics	Type of user test performed	Users involved	User test objective	User test details	Analysis of results
UX heuristics for national park websites [11]	Co-discovery (complemented with a survey)	Sixteen representative users (undergraduate students of tourism, divided into 8 groups of 2 members in each one)	Validate that the problems identified in the heuristic evaluation are perceived as real problems for the users	Yellowstone National Park website was used in the test. The test was designed based on the usability problems identified only by the control group, and rated with a severity greater than 2 and/ or a criticality greater than 4	For each task, the related usability problems and the percentage of fulfillment were detected. Two open questions were asked to identify if the elements considered important for the users are evaluated by the set of heuristics
	Focus group	Sixteen representative users (undergraduate students of tourism, divided into 2 groups of 8 members in each one)	Validate whether the set of heuristics evaluates the features considered relevant by users	The participants discussed the features that a national park website should have. The focus groups lasted approximately 15 min.	The comments obtained were analyzed. Each comment included the associated heuristic that evaluates each comment (if it exists).
Usability heuristics for grid computing applications [22]	Usability test	Five users	Validate whether evaluators using heuristics subjectively ignored the problems	The test was focused on the 6 usability problems identified only by the control group.	Detect if the problems are in fact perceived as real problems by the users.

the set of proposed heuristics via user tests.

- ***Output:⑰User tests results: users'perceptions***

The researcher conducts the user test and interprets the results obtained. It is important to analyze both the quantitative and qualitative results. For instance, to develop heuristics for national park websites, the researchers validated the heuristics through co-discovery and focus group results (see Table 13). The researchers analyzed the quantitative and qualitative data.

For the co-discovery tests, the researchers analyzed the percentage of task fulfillment (quantitative results, Table 14) and user comments (qualitative results, Table 15). Based on the results, the investigators identified that it was not necessary to create new heuristics since all user comments were covered by the heuristics.

During the focus group, the researchers analyzed all relevant topics related to the UX in the national park websites (see Table 13). Then, the investigators checked whether each topic was covered by a related heuristic, and based on the results, they determined that all topics were covered by at least one heuristic. This finding means that the set of heuristics for national park websites evaluates the features considered relevant for the users. In addition, new checklist items were added to the two heuristics (HPN3 and HPN7) based on the users' comments.

4.1.8. Step 8: refinement stage

• ***Inputs: ⑮ heuristic evaluation results: effectiveness of heuristics; ⑯ expert judgment results (survey): utility, clarity, ease of use, need for checklist and comments regarding each heuristic; and ⑰ user tests results: users' perceptions***. To perform step 8, the researcher needs at least one of the 3 inputs proposed. Depending on the validations made in step 7, the researcher can have input ⑮, ⑯ and/or ⑰. Importantly, at least one input is available to perform step 8. For instance, to develop heuristics for the national park websites, the researchers used inputs ⑮ and ⑯ in the first iteration and input ⑰ in the second iteration.

In the refinement stage, the researcher refines the new set of proposed heuristics based on the results obtained in the validation stage. He/she defines the heuristics to be created, refined and/or deleted.

In addition, he/she decides if it is necessary to iterate and apply some stages once again. We recommend performing at least two iterations to validate the set of heuristics twice, performing different experiments and using different case studies.

• ***Output: ⑱ refining document: (1) what heuristics to create, refine and/or delete, why, and how to do it; (2) what steps to repeat***. The researcher analyzes the results using the inputs and determines which changes should be made to the heuristics to improve them. To document the refinements, the researcher can use a table including the ID of the heuristic, its name, the problem detected, and how to solve it. More details on how to document the information can be reviewed in [11].

To develop heuristics for the national park websites, the researchers documented the following refinements: (1) make changes in the definition of some heuristics; (2) add more detail to the specification of the heuristics (more items in the template); and (3) eliminate 5 heuristics of the preliminary set and create 1 heuristic (obtaining a final set of 14 refined and validated heuristics in the second iteration) [11]. In addition, these investigators defined what stages to repeat to improve the set of heuristics. The researchers decided to repeat steps 3, 4, 5, 6, 7, and 8 (see Section 4.2).

In [30], to develop heuristics for smartphones, the authors performed the following refinements: (1) implemented changes in the definition of some heuristics; (2) added more detail to the specification of the heuristics (more items in the template); (3) added new heuristics to cover specific aspects of smartphones that had not been considered at the beginning; and (4) changed the ID of the heuristics [30,47,51].

On the other hand, to develop cultural-oriented usability heuristics [28], authors performed the following refinements: (1) completed the specification of each heuristic by adding checklists [52,53]); (2)

Table 14
Co-discovery test quantitative results.

Task	Usability problem related	Percentage of task fulfillment	Description
Find an event at Mammoth Hot Springs	- The website contains static, very theoretical and irrelevant information. - The calendar icon is not very representative	37.5% (3 of 8 groups)	- 4 groups did not recognize the icon for direct access to the calendar - 3 groups navigated through another section since they could not find the information they were looking for
Search the audio of winter wolves in the audio section	It is confusing to use the search engine (confusing options).	87.5% (7 of 8 groups)	- 7 of 8 groups were able to correctly perform the task - 1 of 8 groups could not use the search engine correctly

Table 15
Co-discovery test qualitative results.

Question	Comment	Number of users who provided comments	Heuristic associated
Q1: What did you like most about the website?	The multimedia resources (images, videos and audios)	10 of 16 users (62.5%)	NPH2 (Multimedia resources)
	The complete information of the park (activities, parks, animals, seasons, and so on)	8 of 16 users (50%)	NPH3 (Information of interest)
	The website design	6 of 16 users (37.5%)	NPH11(Aesthetic and minimalist design)
Q2: What did you like the least about the website?	The difficulty to find specific information (events, prices, itineraries, and so on)	6 of 16 users (37.5%)	NPH3 (Information of interest)
	The lack of an option to change the language.	7 of 16 users (43.75%)	NPH4 (Match between the system and the real world)

changed the specification of the heuristics (ID changes, definition improvements); and (3) eliminated a heuristic from the preliminary set (13 heuristics, [52,53]), obtaining a final set of 12 refined and validated heuristics [28].

4.2. Iterating the methodology stages

It is possible to apply a methodology that follows the 8 stages and complete the entire process without iterations. However, we recommend applying the methodology iteratively and performing at least two iterations to refine and improve the set of heuristics. Satisfactory results are difficult to obtain in a single iteration. To start the second iteration, output ⑱ is used. From the second iteration, not all inputs are mandatory in the stages. The inputs will depend on the activities carried out and the outputs obtained. We recommend using all stages in the first iteration (eventually, step 2 may be optional). Subsequent iterations include fewer stages because they are oriented toward validation and refinement.

On the other hand, we recommend using more than one method to validate and compare the results that are obtained by different methods. We recommend always validating the set of heuristics through heuristic evaluation. The choice of methods will depend on the available resources (e.g., evaluators, experts, and time). The researcher must choose the most appropriate method for the validation (heuristic evaluation, expert judgment, user tests, surveys, focus group, etc.).

The researcher can end the process of creating heuristics if he/she considers that after several iterations and validations, the heuristics are well designed. However, it is difficult to say exactly when the process should stop. A set of heuristics can always be improved. In addition, the rapid evolution of technologies/applications can render a set of heuristics obsolete.

Fig. 2 shows the number of iterations performed to develop heuristics for national parks (diagram a), smartphones (diagram b) [30], and cultural aspects (diagram c) [28]. The iteration number is marked in the diagrams as “It. n”.

To develop heuristics for the national park websites, researchers performed two iterations using our methodology (see diagram a, Fig. 2) [11]. These investigators performed all stages in the first iteration except step 2 and repeated the stages from 3 to 8 in the second iteration. In the second iteration, the researchers repeated the process from step 3 since in the refining stage of the first iteration, they detected new specific aspects that were not covered by the set of heuristics. For step 7 (validation stage), it was relevant to perform experiments with experts in the area of application (tourism students) via user tests. The participants had experience in the domain and could provide useful information for the heuristic validation.

To develop heuristics for smartphones [30,47,51], the authors performed five iterations using the methodology proposed by Rusu et al. [43] (see diagram b, Fig. 2). These researchers performed stages from 1 to 4 in the first iteration to specify the new set of heuristics. Iterations 2, 3, 4 and 5 were used to validate and refine the heuristics, and they performed several experiments for the validation (heuristic evaluations and expert opinion). In the last iteration (iteration 5), they specified the final version of the heuristics.

To develop cultural-oriented heuristics [28,52,53], the authors performed four iterations using the methodology proposed by Rusu et al. [42] (see diagram c, Fig. 2). These researchers performed stages from 1 to 4 in the first iteration to specify the new set of heuristics. Iterations 2, 3 and 4 were used to validate and refine the heuristics. These authors performed several experiments for the validation (heuristic evaluations and expert opinion), and in the last iteration (iteration 4), they specified the final version of the heuristics.

5. Conclusions

The large number of sets of usability/UX heuristics that have been developed to date (more than 80, according to [20,21]) reflect the need to create instruments that allow the effective evaluation of specific application domains (and their specific features). We proposed a formal methodology for establishing new sets of heuristics [11] to facilitate the process of developing usability/UX heuristics. This methodology includes 8 stages. Each stage of the methodology has defined inputs and outputs and details the activities to be performed to create the heuristics. The methodology is flexible and provides the researcher with the possibility of deciding which stages to carry out, how many iterations to perform and which stages to repeat, based on the specific application domain and the available resources (evaluators, experts, and time). It is possible to repeat some (or all) stages, omit stages, and decide which inputs to use in each iteration.

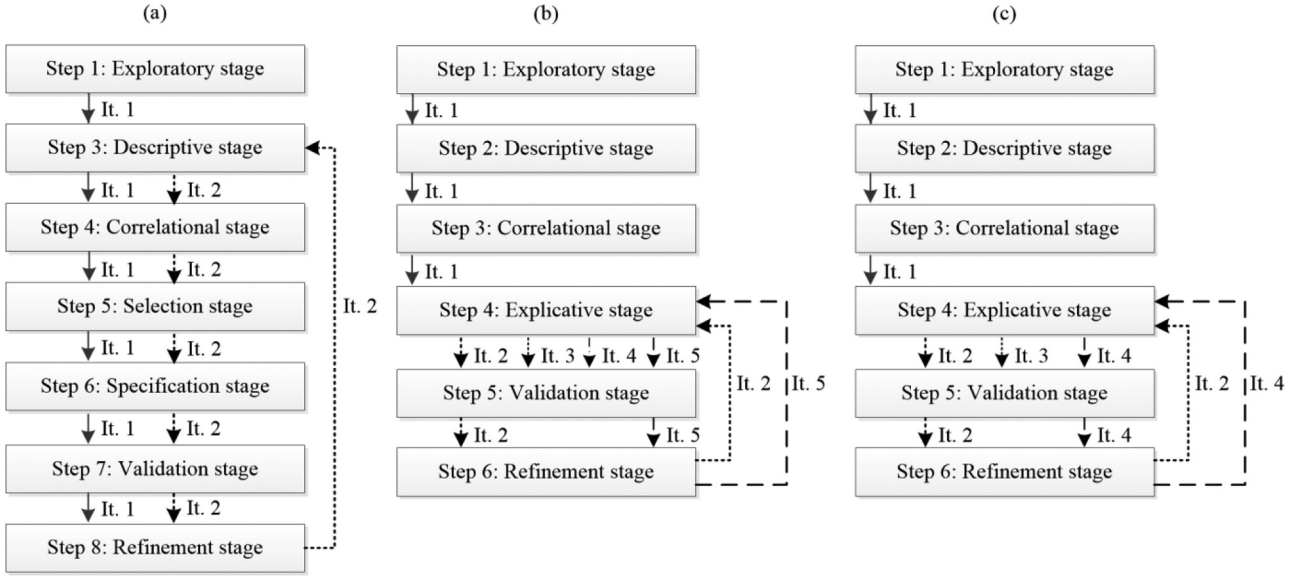

Fig. 2. Iterations performed to develop new sets of heuristics. (a) UX heuristics for national park websites (using the methodology proposed by Quiñones et al. [11]); (b) Usability heuristics for smartphones [30] (using the methodology proposed by Rusu et al. [43]); and (c) Cultural-oriented usability heuristics [28] (using the methodology proposed by Rusu et al. [43]).

As mentioned in [21], developing heuristics should not stop once the heuristics are proposed. It is necessary to perform a robust and rigorous validation in which the effectiveness of the heuristics for specific domains is compared to other heuristics (general and/or specific heuristics). Additionally, standard measures should be adopted to indicate the heuristics' effectiveness. We propose several quantitative and qualitative validation methods in the methodology [11], including precise criteria to evaluate the effectiveness of the heuristics.

To help the experts and/or researchers during the process of applying the methodology, we suggest how to perform each activity, step by step, and how to validate the new heuristics. We also advise how to perform iterations, especially to refine and improve the proposed set. We use several case studies to explain how to apply the methodology. We hope that this detailed explanation will serve as a guide for experts and/or researchers in the creation of new heuristics. We expect that this guide will allow them to see more clearly and concretely what they need to start each stage, how to perform the activities and what they should obtain at the end.

In future work, we plan to fully explain the development of a new set of heuristics for a specific domain by applying the methodology. In addition, we would like to apply the methodology to develop heuristics for attributes other than usability/UX. We are particularly interested in applying the methodology to develop Customer eXperience heuristics [54,55]. If necessary, the methodology will be adapted. We also want to examine how certain stages of the methodology (or possibly all stages) would benefit when using supporting software tools.

Conflicts of interest

None.

Acknowledgments

The authors would like to thank all the participants (experts and researchers) who were involved in the experiments for this study, especially the members of the "UseCV" Research Group in Human-Computer Interaction (School of Informatics Engineering, Pontificia Universidad Católica de Valparaíso – Chile). We would also like to thank the experts Dania Delgado and Daniela Zamora who developed the heuristics for the national park websites for sharing their work. This work was supported by the Universidad de Playa Ancha, Valparaíso, Chile; we would like to thank Prof. Dr. Virginica Rusu and her students.

References

[1] A. Anganes, M.S. Pfaff, J.L. Drury, C.M. O'Toole, The heuristic quality scale, Interact. with Comput. 28 (5) (2016) 584–597.
[2] J. Nielsen, R. Molich, Heuristic evaluation of user interfaces, Proceeding of Conference on Human factors in Computing Systems, SIGCHI '90, 1990, pp. 249–256.
[3] J. Nielsen, "Ten Usability Heuristics", https://www.nngroup.com/articles/tenusability-heuristics/ (accessed 23 November 2018).
[4] C. Rusu, S. Roncagliolo, V. Rusu, C. Collazos, A methodology to establish usability heuristics, Proceedings of the Fourth International Conference on Advances in Computer-Human Interactions, ACHI2011, 2011, pp. 59–62.
[5] D. Van Greunen, A. Yeratziotis, D. Pottas, A three-phase process to develop heuristics, Proceedings of the 13th ZA-WWW Conference, Johannesburg, 2011.
[6] M. Hub, V. Čapková, et al., Heuristic evaluation of usability of public administration portal, in: Narsingh Deo, et al. (Ed.), Applied Computer Science, Proceedings of International Conference on Applied Computer Science, ACS, 2010.
[7] F. Franklin, F. Breyer, J. Kelner, Heurísticas de usabilidade para sistemas colaborativos remotos de realidade aumentada, Proceedings of XVI Symposium on Virtual and Augmented Reality, SVR, 2014, pp. 53–62.
[8] B. Lechner, A. Fruhling, S. Petter, H. Siy, The chicken and the pig: user involvement in developing usability heuristics, Proceedings of the Nineteenth Americas Conference on Information Systems, 2013.
[9] S. Hermawati, G. Lawson, A user-centric methodology to establish usability heuristics for specific domains, Proceedings of the International Conference on Ergonomics & Human Factors, 2015, pp. 80–85.
[10] A.R. Hevner, S.T. March, J. Park, S. Ram, Design science in information systems research, J. MIS Q. 28 (1) (2004) 75–105.
[11] D. Quiñones, C. Rusu, V. Rusu, A methodology to develop usability/user experience heuristics, Comput. Stand. Interfaces 59 (2018) 109–129.
[12] D. Delgado, D. Zamora, Experiencia del Usuario en Sitios Web de Parques Nacionales, Undergraduate Thesis Pontificia Universidad Católica de Valparaíso, Chile, 2017.
[13] V. Roto, M. Lee, K. Pihkala, B. Castro, A. Vermeeren, E. Law, K. Väänänen-Vainio-Mattila, J. Hoonhout, M. Obrist, All About UX, Information for User Experience Professionals (2018), http://www.allaboutux.org (Accessed 15 November.
[14] ISO 9241-210, Ergonomics of Human-system Interaction — Part 210: Humancentred Design for Interactive Systems, International Organization for Standardization, 2010.
[15] T. Jokela, N. Iivari, J. Matero, M. Karukka, The standard of user-centered design and the standard definition of usability: analyzing ISO 13407 against ISO 9241-11, Proceedings of the Latin American Conference on Human-computer interaction, ACM, 2003, pp. 53–60.
[16] P. Morville, "User Experience Design", 2004. http://semanticstudios.com/user_experience_design/ (Accessed 15 November 2018).
[17] M. Revang, "The User Experience Wheel", 2007. http://userexperienceproject.blogspot.com/2007/04/user-experience-wheel.html (Accessed 15 November

2018).
[18] L. Arhippainen, M. Tähti, Empirical evaluation of user experience in two adaptive mobile application prototypes, Proceedings of the 2nd International Conference on Mobile and Ubiquitous Multimedia, 2003, pp. 27–34.
[19] J. Garret, The Elements of User Experience: User-Centered Design for the Web and Beyond, Pearson Education, 2010.
[20] D. Quiñones, C. Rusu, How to develop usability heuristics: a systematic literature review, Comput. Stand. Interfaces 53 (2017) 89–122.
[21] S. Hermawati, G. Lawson, Establishing usability heuristics for heuristics evaluation in a specific domain: is there a consensus? Appl. Ergon. 56 (2016) 34–51.
[22] C. Rusu, S. Roncagliolo, G. Tapia, D. Hayvar, V. Rusu, D. Gorgan, Usability heuristics for grid computing applications, Proceedings of the Fourth International Conference on Advances in Computer-Human Interactions, ACHI, 2011, pp. 53–58.
[23] J. Nielsen, "Usability 101: iIntroduction to Usabilityusability", 2012. https://www.nngroup.com/articles/usability-101-introduction-to-usability/. (Accessed 15 November 2018).
[24] L. Masip, M. Oliva, T. Granollers, User experience specification through quality attributes, Hum.-Comput. Interact. – Interact (2011) 656–660.
[25] N. Aguirre, Experiencia de Usuario en Museos Virtuales, Undergraduate Thesis Pontificia Universidad Católica de Valparaíso, Chile, 2015.
[26] F. Sanz, R. Gálvez, C. Rusu, S. Roncagliolo, V. Rusu, C.A. Collazos, J.P. Cofré, A. Campos, D. Quiñones, A set of usability heuristics and design recommendations for u-learning applications, Information Technology: New Generations, Springer International Publishing, 2016, pp. 983–993.
[27] J.P. Cofré, Usabilidad en u-Learning, Master Thesis Pontificia Universidad Católica de Valparaíso, Chile, 2013.
[28] J. Díaz, C. Rusu, C.A. Collazos, Experimental validation of a set of cultural-oriented usability heuristics: e-Commerce websites evaluation, Comput. Stand. Interfaces 50 (2017) 160–178.
[29] R. Meszaros, I. Lagzi, Z. Barcza, G. Gelybo, GreenView Application Specifications, Eotvos Lorand University, Hungary, 2008.
[30] R. Inostroza, C. Rusu, S. Roncagliolo, V. Rusu, C.A. Collazos, Developing SMASH: a set of SMArtphone's uSability Heuristics, Comput. Stand. Interfaces 43 (2016) 40–52.
[31] A. Campos, C. Rusu, S. Roncagliolo, F. Sanz, R. Gálvez, D. Quiñones, Usability heuristics and design recommendations for driving simulators, Information Technology: New Generations, Springer International Publishing, 2016, pp. 1287–1290.
[32] D. Quiñones, C. Rusu, S. Roncagliolo, Redefining usability heuristics for transactional web applications, Information Technology: New Generations (ITNG), 2014 11th International Conference on, IEEE, 2014, pp. 260–265.
[33] D. Quiñones, C. Rusu, S. Roncagliolo, V. Rusu, C.A. Collazos, Developing usability heuristics: a formal or informal process? IEEE Latin Am. Trans. 14 (7) (2016) 3400–3409.
[34] Real Academia Española, "Definición de parque nacional", 2014. https://dle.rae.es/?id=RyGZA0Z (Accessed 11 March, 2019).
[35] L. Bonastre, T. Granollers, A set of heuristics for user experience evaluation in e-Commerce websites, 7th International Conference on Advances in Computer-Human Interactions, IARIA, 2014, pp. 27–34.
[36] G.H.Hofstede Hofstede, M. Minkov, Cultures and Organizations: Software of the Mind, McGraw-Hill Professional, 2001, pp. 1–29.
[37] L. Smith, T. Dunckley, S. French, Minocha, Y. Chang, A process model for developing usable cross-cultural websites, Interact. Comput. 16 (1) (2004) 63–91.
[38] K.a Huggins, B.B. Holloway, D.W. White, Cross-cultural effects in E-retailing: the moderating role of cultural confinement in differentiating Mexican from non-Mexican Hispanic consumers, J. Bus. Res. 66 (3) (2013) 321–327.
[39] L. Sangwon, R. Koubekb, The effects of usability and web design attributes on user preference for e-Commerce web sites, Comput. Ind. 61 (4) (2010) 329–341.
[40] F. Paz, F.A. Paz, J.A. Pow-Sang, L. Collantes, Usability heuristics for transactional web sites, Proceedings of the 11th International Conference on Information Technology: New Generations, 2014, pp. 627–628.
[41] Korhonen, E.M.I. Koivisto, Playability heuristics for mobile games, Proceedings of Mobile Human-Computer Interactions, 2006, pp. 9–16.
[42] M. Omar, A. Jafar, Heuristics evaluation in computer games, Proceedings of Information Retrieval & Knowledge Management, 2010, pp. 188–193.
[43] C. Rusu, S. Roncagliolo, V. Rusu, C.A. Collazos, A methodology to establish usability heuristics, Proceedings of the Fourth International Conference on Advances in Computer-Human Interactions, ACHI, 2011, pp. 59–62.
[44] Nielsen, "Severity ratings for usability problems", 1995. https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/ (Accessed 20 November 2018).
[45] D.L. Moody, Dealing With Complexity: A Practical Method for Representing Large Entity Relationship Models, University of Melbourne, Department of Information Systems, 2001.
[46] N.C. Fernández, Un Procedimiento de Medición de Tamaño Funcional para Especificaciones de Requisitos (2006) Doctoral Thesis.
[47] R. Inostroza, C. Rusu, S. Roncagliolo, V. Rusu, Usability heuristics for touchscreen-based mobile devices: update, Proceedings of the 2013 Chilean Conference on Human-Computer Interaction, ACM, 2013, pp. 24–29.
[48] C. Rusu, V. Rusu, S. Roncagliolo, D. Quiñones, V.Z. Rusu, H.M. Fardoun, D.M. Alghazzawi, C.A. Collazos, Usability heuristics: reinventing the wheel? Proceedings of the International Conference on Social Computing and Social Media, Springer International Publishing, 2016, pp. 59–70.
[49] V. Rusu, C. Rusu, D. Quiñones, S. Roncagliolo, C.A. Collazos, What happens when evaluating social media's usability? Proceedings of the International Conference on Social Computing and Social Media, Springer, 2017, pp. 117–126.
[50] V. Rusu, C. Rusu, D. Guzmán, S. Roncagliolo, D. Quiñones, Online travel agencies as social media: analyzing customers' opinions, Proceedings of the International Conference on Social Computing and Social Media, Springer, 2017, pp. 200–209.
[51] R. Inostroza, C. Rusu, S. Roncagliolo, C. Jiménez, V. Rusu, Usability heuristics for touchscreen-based mobile devices, Proceedings of the 9th International Conference on Information Technology: New Generations, 2012, pp. 662–667.
[52] C.Rusu Díaz, J.A. Pow-Sang, S. Roncagliolo, A cultural-oriented usability heuristics proposal, Proceedings of the 2013 Chilean Conference on Human-Computer Interaction, ACM, 2013, pp. 82–87.
[53] J. Diaz, C. Rusu, A. Pow-Sang, S. Roncagliolo, Una Propuesta de Heurísticas de Usabilidad Orientados a Aspectos Culturales, 8CCC – Octavo Congreso Colombiano de Computación (2013).
[54] C. Bascur, C. Rusu, D. Quiñones, User as customer: touchpoints and journey map, International Conference on Human Systems Engineering and Design: Future Trends and Applications, Springer, Cham, 2018, pp. 117–122.
[55] V. Rusu, C. Rusu, F. Botella, D. Quiñones, Customer eXperience: is this the ultimate eXperience? Proceedings of the XIX International Conference on Human-Computer Interaction, paper No. 21, ACM, 2018.

Daniela Quiñones was born in Viña del Mar, Chile in 1989. She received the M.S. degree in Computer Science from the Pontificia Universidad Católica de Valparaíso (PUCV), Chile, in 2016 and the PhD degree in Informatics Engineering from the same university in 2018. She is associate professor at PUCV. She has worked in several software companies. Since 2018, she has been dedicated to academia. Her research interests have focused on Human–Computer Interaction (HCI), Customer eXperience, User eXperience, Usability, and Usability Engineering. Dr. Quiñones is member of the "UseCV" Research Group in HCI at PUCV, Chile and member of ACM-SIGCHI.

Cristian Rusu was born in Dej, Romania in 1960. He received the PhD degree in Applied Informatics from the Technical University of Cluj-Napoca, Romania, in 2002. He is full professor at Pontificia Universidad Católica de Valparaíso (PUCV) in Chile. He has worked in several software companies and research institutes in Romania and Chile. Since 1999, he has been fully dedicated to academia. He serves on editorial boards and conference program committees and disseminates User eXperience importance in Chile and Latin America. His research interests are focused on Human–Computer Interaction (HCI), User eXperience, Usability, Customer eXperience, and Service Science. As the former Chair of the Chilean ACMSIGCHI Chapter, he is also the head of the "UseCV" Research Group in HCI at PUCV.