Advancing Sociotechnical-Pedagogical Heuristics for the Usability Evaluation of Online Courses for Adult Learners

Isa Jahnke
Nathan Riedel
Kanupriya Singh
Joi Moore
University of Missouri-Columbia, USA

Abstract
Online courses often include interface designs that do not support a positive learner experience. Literature shows a variety of heuristics to detect issues of online courses. While heuristic-based inspection of usability is a dominant method for evaluating digital systems, these methods cannot be easily transferred to online courses. To close this gap, we identified an initial set of social, technical, and pedagogical related items (STP) heuristics based on literature. Next, we analyzed this set using empirical data from two online courses. In total, we analyzed 195 problems with the goal to substantiate a final set of 14 STP heuristics. This new set allows for efficiently evaluating online courses by supporting evaluators and instructional designers in uncovering the most crucial issues and improving the learner experience. Finally, based on this work, we discuss a definition of learner experience for the emerging field of learner experience design and research.

Keywords: online courses, evaluation, heuristics, online learning, adults, sociotechnical-pedagogical usability

Jahnke, I., Riedel, N., Singh, K. & Moore, J., (2021). Advancing sociotechnical-Pedagogical Heuristics for useability evaluation of online courses for adult learners. *Online Learning, 25*(4), 416-439. DOI: 10.24059/olj.v25i4.2439

Online courses are built with technology, such as Learning Management Systems (LMS) and course authoring tools, that include basic templates that allow some flexibility with the design of the course. The ease of use of such systems is important for supporting a positive experience for the learner, and Nielsen's usability heuristics (Nielsen, 1994) are helpful for understanding the usability of a system and its efficiency, error frequency, and error severity (Botella, Rusu, Rusu, & Quiñones, 2018; Khajouei, Gohari, & Mirzaee, 2018; Sauro, 2014; Stone, Jarrett, Woodroffe, & Minocha, 2005). These methods aim to improve the user experience, which can lead to better engagement with the content. This is important as Demmans Epp, Phirangee, Hewitt, & Perfetti (2020) show that the quality of the system design and course type (student-centered vs. teaching-centered) impact student behavior, experiences, and learning outcomes.

However, as shown by Nokelainen (2006) pedagogical usability is not sufficiently addressed when evaluating online course systems. Moore, Dickson-Deane, & Liu (2014) argue that the traditional technological usability evaluation is not sufficient, and that pedagogical usability is more relevant for learning environments. Other researchers confirm the importance of pedagogical usability (e.g., Horila, Nokelainen, Syvänen, & Överlund, 2002; Lim & Lee, 2007; Quinn, 1996; Reeves, 1994; Silius & Tervakari, 2003).

In addition, a factor that is overlooked and often not considered for the evaluation of online courses is the social dimension. Learning is a social effort, and meaningful online learning is embedded into social group activities (Jahnke, 2015). Learning is dependent on social relations with teachers and peers, as some researchers express the need for humanizing the online space (Jahnke, 2015). Social interactions and social roles are equally important to foster human-centered learning processes. Garrison, Anderson, and Archer's (2003) framework of social, cognitive, and teaching presence indicates the importance of the social dimension as it shows how discussion boards and chats support direct replies and foster learners' interaction.
From this perspective, we propose to evaluate the quality of online courses through the lens of the three dimensions of social, technological, and pedagogical usability. To empirically study an advanced set of sociotechnical-pedagogical (STP) usability heuristics, we examined online courses. We first predefined a set of STP heuristics grounded in literature, then applied them to online courses. This paper presents the results.

The research question was: To what extent is it possible to develop a concise set of sociotechnical-pedagogical heuristics (STP heuristics), and what is the quality of the heuristics when applying them to online courses?

Review of Related Work

Usability of online courses is measured by the learner's interaction with the learning management system (LMS) that also includes the course organization, material presentation, sociability, and other elements of the LMS. Studies have shown the importance of evaluating the usability of online courses; however, studies have used different perspectives (Dringus & Cohen, 2005; Mayer, 2002; Reeves, 1994). There exists a gap between the social, technological, and pedagogical usability aspects of online courses. According to Guo, Kim, and Rubin (2014), some perspectives include the technical or the pedagogical approach to course usability. Nokelainen (2006) focused on the social and pedagogical aspects of online courses.

Heuristics in General

Heuristics are used to evaluate the user-friendliness and usability level of digital systems. A heuristic evaluation uses a set of items and applies them to a certain system or technology with the goal of detecting potential issues for the purpose of improving the technology and the user experience (e.g., Nielsen, 1994). A prominent heuristic tool developed by Nielsen (1994) includes a set of 10 heuristics to guide designers in detecting technological usability problems in systems.

Recently, in the field of online courses, effective design indicators have been developed. Design and evaluation instruments, such as Quality Matters, digital didactical designs, and the 12 principles of multimedia learning (Quality Matters, 2018; Jahnke, 2015; Mayer, 2002) work as a rule of thumb when creating online courses. For example, Quality Matters contains eight general items, each broken down into more detailed items, that guide education professionals in aligning learning objectives, activities, and learner support (Quality Matters, 2018). Such guidelines allow for quick, high-quality course design.

Technological, Pedagogical, and Social Usability Heuristics

Nielsen and Loranger (2006) define usability as “how quickly people can learn to use something, how efficient they are while using it, how memorable it is, how error-prone it is, and how much users like using it. If people can’t or won’t use a feature, it might as well not exist” (p. xvi). Usability focuses on the optimization of user interaction with the interface to enable the user to perform typical tasks. It also includes the evaluation of aesthetic features to support a positive user experience with the system. In this study, we refer to this kind of usability of the interface interaction as *technological usability***.** For example, in online courses, learners interact with the interface features of a learning management system, such as navigating to resources, viewing grades, creating a post in the discussion board, submitting assignments, and so forth. The usability of the system can affect the learner experience and learning performance with the online course.

However, interface interaction (technological usability) alone may not explain the entire learner experience. The qualities of technology-related usability are not sufficient to guarantee that an online course leads to a positive learning experience for learners. The pedagogical and social aspects related to the design of the learning process, communication among students and teachers, purpose of learning, content arrangement, and learning strategies applied, all support the achievement of learning objectives and create meaningful learning experiences for learners (Jahnke, 2015; Lim & Lee, 2007). A concise set of social and pedagogical usability heuristics would unpack such aspects.

Social usability in this paper comprises the learner’s activities with other learners, such as computer-mediated communication with peers or interactions with the tools of the online course. Social usability focuses on human-human interactions supported by technology (Preece, 2001). Jahnke et al. (2005) showed the relevance of formal and informal role dynamics and how they affect learning or interactions, e.g., having access or not to certain tools or files in the course, or role changes during a certain time. Their study indicated that the evaluation of the LMS tools to support social dimensions of learning technologies has been neglected. Robinson, Sheffield, Phillips, and Moore (2017) found that social interactions in online courses have a positive impact on student perceptions. Similarly, studies of social usability in online courses have found that level of interactivity, social presence, and student characteristics in online courses significantly impact the online learning experience for students (Chen, Chang, Ouyang, & Zhou, 2018; Kaufmann, Sellnow, & Frisby, 2016; Martin & Bolliger, 2018; Orcutt & Dringus, 2017).

According to Silius and Tervakari (2003), *pedagogical usability* refers to whether the tools, content, interface, and tasks in an online learning environment support a variety of learners in achieving learning goals and objectives. Though pedagogical usability is less frequently studied than technical usability (Nokelainen, 2006), there exist pedagogical usability frameworks and heuristic checklists for evaluating online courses or web-based learning (Albion, 1999; Horila et al., 2002; Lim & Lee, 2007; Moore et al., 2014; Nokelainen, 2006; Quinn, 1996; Reeves, 1994; Silius & Tervakari, 2003; Squires & Preece, 1999). In their recent work, Yousef, Chatti, Schroeder, and Wosnitza (2018) demonstrate that effective learning design can improve pedagogical usability and make online courses more motivating for learners.

For this work we refer to Jahnke, Schmidt, Pham, & Singh (2020), who defined a conceptual framework of sociotechnical-pedagogical usability. Basically, we define sociotechnical-pedagogical usability with three dimensions that include the following elements:

Social: teacher or learner communication, collaboration or group learning, human interaction by means of digital tools, social presence, social roles/relationships
Technical: usability related to technological issues
Pedagogical: teaching or learning goals, student activities, assessment

To develop a new set of sociotechnical-pedagogical usability heuristics for online courses, we first applied a literature review before we tested the STP heuristics empirically (see Method section). For the literature review, thirty articles were reviewed in total. The research team contributed to the collection of articles. In general, articles were selected if they included the key words "online course usability," "online course recommendations," or "online course design principles." In detail, articles about designing, evaluating, or improving online courses with a focus on social aspects of technology, use in education, or just pedagogy were selected. In addition, we looked at articles that consisted of different principles, heuristics, and guidelines ranging from system usability to pedagogical theories. Table 1 lists all 30 publications. The 30 articles from the literature review have been used to derive items for the development of STP heuristics.

Table 1
Breakdown of STP Heuristics Derived from Literature Review

Source (alphabetical order)	**Year**	**STP**	**Items derived from literature**	**No. of items**
Benson et al.	2002	P, T	Technology interactions, learning products adhere to widely recognized standards for technology/ software interactions.	17
Bloom	1956	P	Objectives are developed based on Bloom's taxonomy.	1
Boyle	1997	T, P	Give learners controls (e.g., pause, go back, go forward, skip) to allow them to access the video at their own pace.	1
Chao, Saj, & Tessier	2006	P	Language use is consistent throughout the course.	4
Clement	1985	T	When presenting one topic/idea, follow the "rule of seven" guideline: present a maximum of seven pieces of content at a time.	1

Das	2012	P	Syllabus contains information regarding drop/withdraw/return policy.	1
Douglas	2017	T	Pages and sections mentioned in the instructions or throughout the course should include a link to provide shortcuts for efficient navigation.	1
Dringus & Cohen	2005	P	Content elements are presented in a logical sequence.	18
Fink	2012	P	Syllabus provides titles of assignments and relevant points.	1
Guo et al.	2014	P, T	Videos should display the instructor's talking head at opportune times.	4
Jahnke	2015	P, S	Learning activities are active and facilitate engagement via learner-content, learner-learner, and learner-instructor interactions.	21
Lenzner, Schnotz, & Müller	2013	P	If used, images should be relevant to learning content and enhance the knowledge acquisition.	1
Mayer	2002	P, S	People learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen.	8
Moore et al.	2014	P	Content can be organized using hierarchical classification.	12
Nielsen	1994		Users should not have to wonder whether different words, situations, or actions mean the same thing.	4
Nielsen	2004	T	Underlines are only used to indicate working links to relevant sections.	1
Nokelainen	2006	P, S	Authentic stories, anecdotes, emotion, or human conflict are used to engage learners and show real-world relevance when appropriate.	2
OLC	2018	P	Syllabus communicates expectations for students and discussion participation.	12
Obsidian Learning	2017	P, T	Keep videos short. However, video length should be governed by the nature and complexity of the content.	1
Quality Matters	2018	P, S	Information and instructions are provided regarding how the tools support the learning objectives or competencies.	35
Reeves et al.	2002	P, S	The interactivity with technology has meaningful learning purposes.	1
Reeves	1994	P	The objectives/goals of the course and each module are present so learners know what objectives/goals they can achieve.	4
Safie	2007	T	Technology is compatible with all devices.	1
Schade	2014	T	Users must be able to interact with videos as they often do in their daily lives, such as watching in full view or playing backward or forward.	1
Sims, Dobbs, & Hand	2002	P	The manner of submission for assignments/assessments is clear.	2
Stein & Graham	2014	P	Materials consistently indicate when activities or assessments take place on site versus online.	1
Stone et al.	2005	T	If something is important for the user, it should be placed in a prominent position.	5
van der Meij & van der Meij	2013	P, S	Draw attention to the interconnection of user actions and system reactions.	14
Van Merriënboer, Kirschner, & Kester	2003	P	Introduce new concepts by showing their use in context. In other words, knowledge is presented at the point when the user needs that information to perform the task.	1
Xavier University	2018	P, S	Syllabus contains information regarding the course summary or the main parts of the course.	13
Zhang, Zhou, Briggs, & Nunamaker	2006	P	Interactive video is preferred over non-interactive video.	1

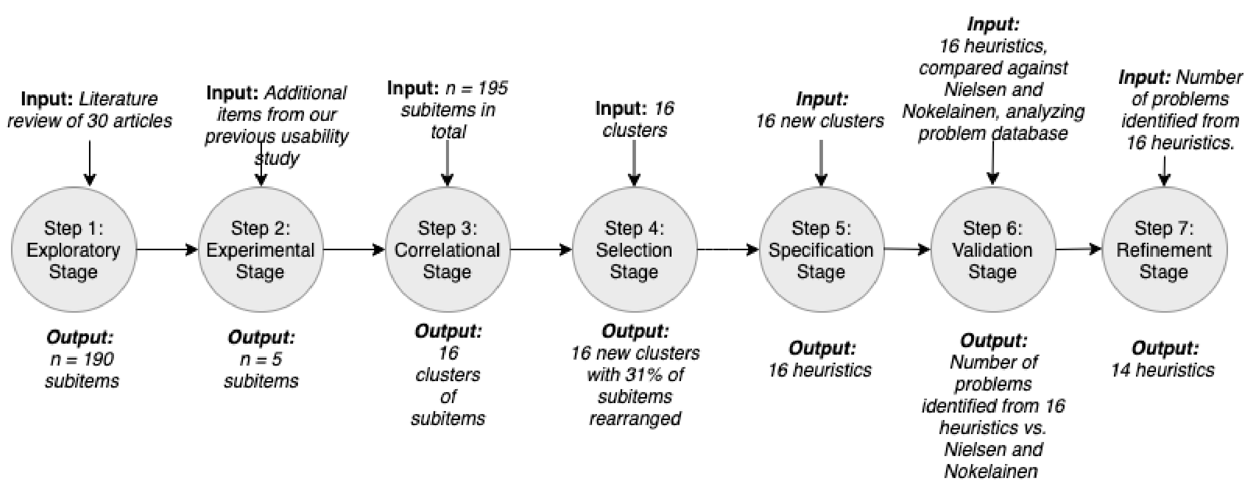
Research team	Unpublished	T, P	Additional items are from previous user experience studies related to technology-enhanced learning (not found in literature): provide hierarchy of content, provide same page title, ensure page title and page content match, describe acronyms or abbreviations, and add navigation instructions.	5
Total				**195**

Note. *N* = 195. Items are principles or guidelines that were integrated into a new set of sociotechnical-pedagogical heuristics for online course usability evaluation (see Method section).

Method

The goal of the study was to develop and test key sociotechnical-pedagogical heuristics for evaluating and detecting issues in online courses. We applied the heuristic development methodology guided by Quiñones, Rusu, and Rusu (2018). They provide a roadmap to ensure quality, reliability, and validity when developing new heuristics. The final heuristic development framework of Quiñones et al. (2018) consists of eight steps. However, Quiñones et al. (2018) also stress that some steps may be omitted if they are unnecessary based on context or that some steps may overlap as they may need to occur simultaneously. Our method consisted of seven steps, as outlined in Figure 1. We describe the process and methods of each step in the following sections. This process led us to a final set of 14 heuristics that are described in the Results section.

Figure 1
Steps of developing and testing a new set of STP heuristics



Note. Steps are adapted from Quiñones et al. (2018)

In total, 13 research team members were involved in different phases of the project: three master students or interns in the study program of learning design & technologies, nine doctoral students of information science and learning technologies, trained in usability evaluation, and one expert; see Appendix A for details.

Steps 1 and 2: Exploratory and Experimental Stages

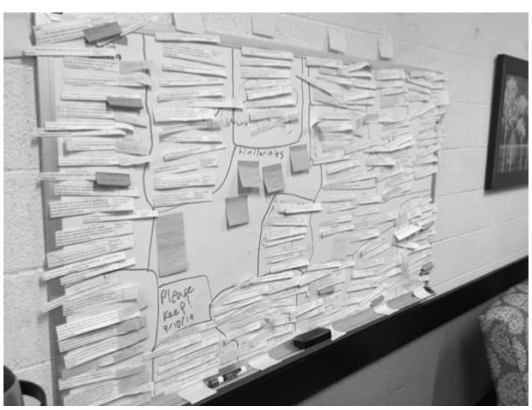
In Step 1, we collected 190 items from literature based on 30 articles (see Table 1). The 190 items from literature were collected by searching repeated patterns of social, technological,

and pedagogical principles in literature of online learning environments. Five members of the research team collected the literature. The members chose articles based on several criteria. Members searched for established instructional design guidelines (e.g., Quality Matters) and articles with outcomes that provided recommendations for instructional design (e.g., Nokelainen, 2006) or for the creation of instructional content (e.g., Fink, 2012). In addition, they searched for articles that focused specifically on the use of technology in education (Stein & Graham, 2014). In summary, criteria for the literature search included design guidelines or principles for designing online learning from social, pedagogical, or technological views. All articles were compiled in a shared, cloud-based document to ensure that no articles were duplicated.
Step 2 in this study followed the experimental stage of Quiñones et al. (2018), which recommends adding additional items identified via specific features of the application, detected usability problems, and problems with existing heuristics. We added five items from our user experience tests from technology-enhanced learning environments that were not found in existing literature. In total, there were 195 items because of these two steps.

Steps 3–5: Correlational, Selection, and Specification Stages

In Step 3 (Correlation Stage), the 195 items were analyzed for correlation. The research team took the 195 items and matched them to similar items. In detail, after collecting a total of 195 items from the literature, printed versions of the items were posted on a whiteboard (see Figure 2). Then, team members collaboratively began placing items that addressed similar issues next to one another (e.g., aesthetic guidelines and course material guidelines). In the process of clustering relevant data, the team members began coding similar features as they appeared by naming each category. During this process, categories were consolidated or split based on whether the team members felt a category was too narrow or included too many topics. As similarities began to take shape, the team members suggested names for each category. Any item that did not fit one of the categories or required further details was placed in the center of the wall in a miscellaneous category; these items were discussed later and placed into an existing category or a new one was created. Once all the items were placed into categories, the items were then transferred into a *digital* list of 16 categories.

Figure 2
Initial categories of items displayed on a white board



In Step 4, Quiñones et al. recommend conducting a Selection Stage, in which heuristic developers keep, adapt, and/or discard the heuristics developed in the previous step. Hence, we refined the collection of the 16 categories. In detail, the list with the 16 categories was sent off to a research group member who is an expert in technology-enhanced and online learning design and who was not involved in the clustering. This member then read the categories titles and their descriptions. In cases of non-agreement, she offered new title suggestions and definitions. Approximately 31% of the items were moved or re-organized. The result was a list of 16 partly revised categories.

Finally, in Step 5, research team members, who had two or three years of experience and solid skills in system usability evaluation, were asked to review the categories in order to recommend their own names and any suggestions for moving any items to a new category. If an item was disputed, they discussed its best placement until a unanimous agreement was reached. The team ensured that all miscellaneous items were meticulously discussed and assigned to a category. Each category was then named a *heuristic*. The result of Step 5 was a preliminary set of 16 heuristics, with names, that were iteratively developed bottom-up from coded items (Steps 1–5).

Steps 6 and 7: Validation and Refinement Stages

The next two steps focused on ensuring the quality of the 16 new STP heuristics. In Step 6 of this study, researchers applied two forms of validation methods, which are both recommended by Quiñones et al. (2018). The first validation method is called the expert review, in which the research team members took on the role of evaluators and applied the preliminary heuristics to detect problems in online courses. More specifically, to validate the 16 STP heuristics, we checked them against a problem database that included 144 problems from two online courses. (Details of the database development are in the next section.) Each of the 16 STP heuristics was assigned to the 144 problems identified in the two online courses. More than one heuristic could be applied to each problem. All 144 problems were put in a digital spreadsheet with their assigned heuristics. Three research team members conducted this procedure. Each researcher's set was then analyzed for interrater reliability using a Fleiss's Kappa test. If there was no consent, meaning all three members selected three different heuristics for the same problem, then they met to discuss their decisions. In a few cases, no consensus could be reached (see Results).

In the second validation method, the team compared the new STP heuristics with previously established heuristics sets of Nielsen (1994) and Nokelainen (2006). The Nielsen set was chosen because of its technology-centric heuristics and Nokelainen heuristics focus on pedagogical usability in technology-enhanced learning. According to Quiñones et al. (2018), the purpose of this form of validation is to determine whether the new heuristics are able to diagnose issues not identified by older heuristics.

The result includes a table (see Table 5) with problems identified using either the Nielsen-, the Nokelainen-heuristics, or the new STP set. Based on the results, we were able to refine the 16 heuristics into a final set of 14 STP heuristics. For the refinement stage, we assumed that heuristics assigned less frequently to the 144 problems could be merged. In addition, we applied plausibility and a content view, meaning if two or more heuristics addressed similar problems, they could be merged.

Problem Database

The problem database was developed based on user experience studies for two online courses. These two courses are titled *Master Gardener* and *Fire Service Instructor I*, and both are taught in Canvas. We describe the courses then the problem database.

Master Gardener was a 14-week online course offered by the extension division of a Midwestern university in the United States. The course was offered during Spring 2019 (January to May 2019) and focused on topics related to gardening. The course was designed for adult learners who wanted to advance their knowledge of horticulture and intended to become certified master gardeners. There were 60 to 70 students enrolled of various age ranges, mainly falling within categories of ages from 35 to 39, 40 to 49, and 50 years and older. These learners were largely from rural areas of a Midwestern state in the United States.

Fire Service Instructor I (FRTI-Instructor I) was a seven-week course offered by the extension division of Midwestern university in the United States. The course was designed for firefighters who wanted to pursue professional careers as firefighter instructors. The course was mainly online but had one face-to-face meeting in the first week, in which the instructor described the course process. Enrolled students met for a face-to-face session on the first day of the class from 8 am to 4 pm for class introduction, goals, and objectives of learning. The rest of the seven-week class was offered online. According to the instructor, an adult learner needed a total of 40 hours to complete the course and receive a certificate of completion. The online portion of class comprised of various learning activities, such as assignments, quizzes, and discussions.

Each course underwent a usability study to detect potential issues with the online course. The reports of these two studies were the foundation for developing the problem database. The database was created by utilizing the usability problems discovered in the two online courses. The database began as two online spreadsheets, one for each of the online courses. The Fire Service spreadsheet was developed by using an expert evaluation report that was conducted in May 2019. Each problem from this report was placed in the first spreadsheet of the database. The Master Gardener spreadsheet was created using the results of the interviews with the participants who were enrolled in the course. Problems that emerged from the interviews were placed in the second spreadsheet of the database. The two spreadsheets were then merged. In summary, the database consists of a mix of problems identified by experts and students.

Between the two courses, a total of 144 problems were identified, with 76 problems from the Fire Service Instructor course and 68 from the Master Gardener course. The identified problems ranged across issues. Some issues were related to the objectives and goals of the course while other problems related to the course content. Some problems referred to the system of the courses, such as action buttons or multimedia problems. Additional problems included page layout (e.g., font size) and lack of accessibility (e.g., the course not providing alt text for the pictures). Both spreadsheets together compose the problem database.

The list of all 144 problems can be accessed online at https://sites.google.com/view/stp-heuristics/problem-database. The problems are labeled with FS or MG to identify the course (FS = Fire Service Instructor; MG = Master Gardening).

Results

The 16 heuristics developed from the literature analysis are presented here followed by the refined STP heuristics that were analyzed with two online courses.

Results from Steps 1–2 and Steps 3–5

Overall, 190 literature items and 5 additional items from our previous studies (195 in total) have been used for the development of a new heuristics set. Following the process of steps 3, 4, and 5 as described in the Method section, the correlation (coding and clustering), selection, and specification led to the result of a preliminary set of 16 STP heuristics, as shown in Table 2.

The detailed list of the heuristics with all 195 coded items can be found at https://sites.google.com/view/stp-heuristics/home.

Table 2
The Preliminary 16 Heuristics

H #	Heuristic	STP	Description	# of items	Typical examples
1	Social Presence	S	Refers to social aspects of the course (e.g., communication, social or teacher presence of instructor). It points to potential issues of instructor-student communication, in online discussion boards, or of student roles in teamwork.	9	1.4 The course provides learners with opportunities to access extended feedback from instructors, experts, peers, or others through e-mail or other Internet communications (Benson et al., 2002). 1.6 Instructor plays different roles (e.g., expert, mentor, coach, learning companion) (Jahnke, 2015).
2	(Group) Activities	S, P	Refers to (group) learning activities and assignments within the course. It points to potential issues of quality of learning activities (e.g., assignments), or activities that do not match learning objectives.	15	2.8 The course supports various modes of learning, including group activities (Dringus & Cohen, 2005). 2.11 The purpose of the activities is clearly stated so students understand how they tie into course objectives (Jahnke, 2015).
3	Easy to Use	T	Refers to technological usability. It points to potential issues of accessing course materials or completing activities (e.g., uploading files).	8	3.2 Users should not have to wonder whether different words, situations, or actions mean the same thing (Nielsen, 1994). 3.7 Users are able to edit their own and reply to others' messages in discussion posts (Dringus & Cohen, 2005).
4	Page Layout	T	Refers to the aesthetic design of an online course. It points to potential issues with font size, color, chunking of text (i.e., leaving white space in between sections of text), etc.	28	4.5 No extraneous or irrelevant information, visual noise, or unnecessary styles are present (Moore et al., 2014). 4.24 The design and presentation of information is consistent (e.g., layout, color, text size, text style, font) (Stone et al., 2005).
5	Ecosystem	T	Refers to the broader learning management system's capabilities. It points to potential issues with correct use of the menu function of the LMS or the organization of modules.	17	5.1 If the course includes links to external resources, the links are kept up to date (Benson et al., 2002). 5.4 Frequently used technology tools are easily accessed (Quality Matters, 2018).

6	Navigation	T	Refers to the design of navigation within the online course. It points to potential issues with searching for information, modules, or instructions in the LMS.	8	6.2 Course design enables learners to easily locate where they are within the course (Online Learning Consortium, 2018). 6.4 Related materials and resources are joined by hyperlink (Moore et al., 2014).
7	Functionality	T	Refers to functionality from the view of human-computer interactions. It points to potential issues of lack of feedback students receive from the system, or device compatibility.	10	7.3 The system is designed so that the learner recognizes when and where he/she has made a mistake (Nielsen, 1994). 7.9 Online resources open in new windows (Chao et al., 2006).
8	Accessibility	T	Refers to accessibility rules (e.g., ADA violations). It points to potential issues with accessibility rules (e.g., a lack of accessibility statements or direct links to institutional accessibility policies).	7	8.5 The course provides guidelines and/or Q&A for disabled students to seek technology and/or academic help (Quality Matters, 2018). 8.7 For accessibility, provide a means for the learner to access the text of the narration (van der Meij, 2013).
9	Diverse Material	P	Refers to material being used in the online course. It points to potential issues of having too much material be too similar (e.g., too many videos and no other types of materials), quality level of video narration, or repetitive content.	16	9.2 Tablet drawing tutorials (e.g., Khan-style table drawing tutorials) are more engaging than PowerPoint slide presentations with voice-over (Guo et al., 2014). 9.13 People learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen (Mayer, 2002).
10	Material Organization	P	Refers to how and when materials are arranged within the course. It points to potential issues of material being extraneous to the learning objectives of a module or too much information included per module.	15	10.6 Introduce new concepts by showing their use in context. In other words, knowledge is presented at the point when the user needs that information to perform the task (van Merrienboer, Kirschner & Kester, 2003). 10.7 Information and instructions are provided regarding how the tools support the learning objectives or competencies (Quality Matters, 2018).

11	Material Delivery	P	Refers to how material is presented to students (e.g., introducing concepts or providing questions for learners to consider). It points to potential issues of material not being properly scaffolded (i.e., introduced too early or late in the course) or being randomly added to modules.	7	11.2 One topic or idea is introduced at one time (Moore et al., 2014). 11.7 Content elements are presented in a logical sequence (Dringus & Cohen, 2005).
12	Material Quality, Interactive Material	P	Refers to the quality of material used (e.g., quality of videos, textbooks, open access). It points to potential issues with how materials allow users to interact with the content or how much cognitive load the material requires from the learner, or with how up to date material is.	13	12.1 "There is no extraneous processing in using materials, resources, and multimedia" (Moore et al, 2014). 12.5 "Interactive video is more preferred than non-interactive video" (Zhang et al., 2006).
13	Assessment	P	Refers to forms of assessment in the course. It points to potential issues with quality of assessments or timeframe for feedback on assessments.	17	13.2 Activities and assessments are adequate and reasonable for the course duration (Xavier University, 2018). 13.5 Ongoing assessments are conducted to verify the learner's readiness for the next lesson (Jahnke, 2015).
14	Syllabus	P	Refers to the written syllabus being easy to find and having meaningful content. It points to potential issues of not finding the syllabus or it lacking relevant information.	7	14.2 Syllabus contains information regarding instructor presence and response time for assignments (Xavier University, 2018). 14.6 Course overview and/or introduction, includes pre-requisite knowledge in the discipline and/or any required competencies that are required for the successful completion of the course (Quality Matters, 2018).
15	Teaching/ Learning Goals	P	Refers to learning goals/objectives. It points to potential issues with quality of learning objectives and how they will be measured or conveyed to learners.	7	15.2. All learning objectives are stated clearly, written from the students' perspective, and prominently located in the course (Quality Matters, 2018; Jahnke, 2015). 15.5. The objectives/goals of the course and each module are present so learners know what objectives/goals they can achieve (Reeves 1994).
16	Guidance	P	Refers to course information to guide students.	11	16.3 Information on how to get started is present and stands

		It points to potential issues of not informing students on topics such as where to find information or how to access help (i.e., technical, or educational).		out in the home page (Xavier University, 2018). 16.5 All help and documentation are written clearly and succinctly (Benson et al., 2002).

Note. Numbers such as 1.4, 2.8 are examples of the coded items and the full list is available at https://sites.google.com/view/stp-heuristics/home

Results from Steps 6–7 (The Empirical Study)

To test the thoroughness of the preliminary 16 STP heuristics, the research team assigned a heuristic to each of the 144 problems identified from the previous usability studies of two online courses, as described in the Method section. Problems could be assigned to multiple heuristics. The three heuristics with the highest level of frequency were Material Delivery, Guidance, and Material Quality, with each being assigned to 18 or 19 problems. Diverse Material had the lowest level of frequency with only 4 problems assigned to it. The research team had difficulty reaching consensus regarding which heuristic to assign to seven of the problems (as indicated in the problem database). In such cases, the three evaluators had assigned three different heuristics while the other problems had a consensus of one or two heuristics. For a more detailed breakdown of problem frequency see Table 3.

Table 3
Problems per Heuristic

Heuristic no.	Heuristic name	STP	Frequency (problems assigned to heuristic)	# of problems assigned to additional heuristic(s)
11	Material Delivery	P	19	12
16	Guidance	P	18	2
12	Material Quality/Interactive Mat.	P	18	2
4	Page Layout	T	17	1
15	Teaching/Learning Goals	P	11	1
6	Navigation	T	10	3
2	Activities	S, P	10	5
7	Functionality	T	9	3
14	Syllabus	P	9	0
8	Accessibility	T	9	0
10	Material Organization	P	8	3
1	Social	S	7	0
13	Assessment	P	6	0
3	Easy to Use	T	6	3
5	Ecosystem	T	5	2
9	Diverse material	P	4	2

Note. *N* = 144 problems assigned to heuristics. A problem can be assigned to more than one heuristic.

The quality test of the 16 heuristics shows that some of the heuristics were assigned to 17–19 problems while other preliminary heuristics were only assigned to 4–8 problems. Based on plausibility, this was an indication that some of the heuristics assigned less frequently could be merged. To merge them also makes sense from content view because they address similar problems. Based on the data, H9 (Diverse Material) and H12 (Material Quality/Interactive Material) were merged. Diverse Material and Material Quality both contain items that could inform one another. Moreover, Diverse Material was only assigned to four problems in total while Material Quality was one of the two heuristics most frequently assigned. Additionally, H10 (Material Organization) and H11 (Material Delivery) were also merged into one heuristic. Combining these four heuristics into two would allow for the list to maintain its integrity but also become more condensed (Step 7; see Table 4). As Nielsen argues that heuristics don't have to be distinct and can partly overlap if they help to detect the problems with the digital system (Nielsen, 1994), there was no need for additional merges. Table 4 shows the final set of 14 STP heuristics.

Table 4
Refined and Final Set of 14 STP Heuristics

No.	Final set of STP heuristics	STP (merged)
1	Social Presence	S
2	(Group) Activities	S, P
3	Easy to Use	T
4	Page Layout	T
5	Ecosystem	T
6	Navigation	T
7	Functionality	T
8	Accessibility	T
9	Diverse Material/Quality	P (#9 and #12)
10	Material Delivery/Organization	P (#10 and #11)
11	Assessment	P
12	Syllabus	P
13	Teaching/Learning Goals	P
14	Guidance	P

In Step 7, we ran checks against two previously established sets of usability heuristics: Nielsen (1994) for technical usability and Nokelainen (2006) for pedagogical usability. There were no existing heuristics for the social dimension. Both sets of heuristics are established heuristics. For example, Nielsen is used in industry and is considered a standard in usability evaluation. Two teams conducted the cross-checking against the new STP heuristics, each taking one of the previously established heuristics, either Nielsen (1994) or Nokelainen (2006).

Team Nielsen was able to identify 129 of the 144 problems. The research team evaluated both the design of the technology (learning management system) as well as the instructions integrated in the technology (pedagogy). For example, making information easily accessible to students refers to both technological and pedagogical design decisions; such a design may impact the ways in which users interact with a system. Fleiss's Kappa was used to determine interrater reliability among the research team and resulted in substantial agreement (62%) when applying Nielsen to the problem database. With the Nielsen heuristics, only 128 of 144 problems would have been found or detected.

Team Nokelainen was able to identify only 90 of the 144 problems. Fleiss's Kappa was used to determine interrater reliability among the researcher team and resulted in moderate agreement (60%). These results show that the new set of 14 STP heuristics do identify more issues than Nokelainen's heuristics and demonstrate the quality of this new set of STP heuristics (see Table 5).

Team STP heuristics was able to identify all 144 problems but had difficulty with 7 problems where no consensus was reached. The final Fleiss' Kappa score on the STP heuristic assignment was substantial with 80% reliability.

In addition, Kappa was also used to determine the severity of the problems found in the two courses (see Table 5). Severity should be tracked alongside, yet independent of, problem frequency to determine which problems require attention over others so as not to frustrate users (Sauro, 2014). Three raters assessed the severity of each problem by assigning the problem a value between one and five, with one being minor in severity and five being major in severity. This test was used to determine if the severity of the problems assigned to the heuristics were similar across the three raters. The final Kappa score on problem severity was substantial (64%). The results of the Fleiss's Kappa suggest the heuristics can accurately identify sociotechnical-pedagogical usability issues with varying severity.

Table 5
Comparison of Previously Established Heuristics and STP Heuristics

Heuristics Set	Problems detected	Problems not detected	Severity of problems detected	Severity of problems not detected
Nielsen	129	15	Level 5 = 34 Level 4 = 48 Level 3 = 28 Level 2 = 11 Level 1 = 1 Undecided = 7	Level 5 = 5 Level 4 = 4 Level 3 = 5 Level 2 = 1 Level 1 = 0 Undecided = 0
Noke-lainen	90	54	Level 5 = 22 Level 4 = 37 Level 3 = 23 Level 2 = 4 Level 1 = 1 Undecided = 3	Level 5 = 15 Level 4 = 17 Level 3 = 10 Level 2 = 8 Level 1 = 0 Undecided= 4
STP	144	0	Level 5 = 39 Level 4 = 52 Level 3 = 33 Level 2 = 12 Level 1 = 1 Undecided = 7	0

Note. N = 144 problems. Severity level based on consensus of 2 out of 3 raters using a severity scale of 1 to 5 (1 = minor, 5 = major) with undecided indicating raters do not agree.

Discussion

The final set of 14 STP heuristics developed through this process are detailed and robust enough to address potential issues in online courses. This study's research question contained two parts. First, to what extent is it possible to develop a concise set of sociotechnical-pedagogical heuristics? The research team was able to develop heuristics that could be condensed to provide a more concise guide for evaluation of or troubleshooting for online courses (Quiñones et al., 2018). Using Quiñones et al. (2018) as a guide, this study has resulted in a set of STP heuristics that can identify a variety of problems including social usability (heuristics 1 and 2), physical design of the course (heuristics 3 through 8), material selection and delivery (heuristics 9 and 10), and pedagogical usability, including assessment and teaching/learning goals or objectives (Heuristics 11 through 14).

Second, what is the accuracy of the heuristics when applying them to online courses? Quiñones et al. provided a method for refining and testing the quality of a new set of heuristics. By using the method, the research team tested the STP heuristics against two control heuristics (i.e., Nielsen and Nokelainen) and demonstrated the ability of the STP heuristics to identify problems that would have been neglected by the control heuristics. As outlined in Quiñones et al. (2018) and Sauro (2014), the new heuristics should exceed the control heuristics in identifying issues both in number and variety of severity levels. The procedures used in this study identified 60 problems, with varying levels of severity, that would not have been identified by a combination of both Nielsen (1994) and Nokelainen (2006). Overall, the STP heuristics were able to identify several issues (see Table 5) that neither Nielsen's nor Nokelainen's heuristics detected. Some examples include instructor self-introduction and social presence (heuristic 1), appropriate placement of course syllabus (heuristic 14), video length (heuristic 9), and page/font formatting (heuristic 4). The thorough validation method used demonstrates both the gaps that exist in current heuristics and the strength of the new STP heuristics.

The STP heuristics were developed from literature and checked against online courses (*Fire Service Instructor* and *Master Gardening*) for adult learners. The heuristics properly addressed the problems discovered in the fire service instructor course. Furthermore, the heuristics were sufficiently assigned and were able to address every problem identified in both the fire service instructor and master gardener courses, with only 16 of 144 problems being assigned to more than one heuristic and none of the problems going unassigned. The new set of STP heuristics developed here can be used for the evaluation of online courses. We assume that the evaluator should be a team of two or three members. Having evaluators who are trained in usability evaluation or who have an instructional design experience may be an advantage; however, further research is needed about the skills of such evaluators.

Third, heuristics have been developed in the field of user experience (UX) for software development and marketing fields. This new set of STP heuristics is an early step in using UX methods in digital learning, which is emerging as a new field of learner experience research (Schmidt et al., 2020). This new field of learner experience (LX), is at the crossroads of UX, learning design, and educational technology. However, there is no common or shared understanding yet of what learner or learning experience is. With this first work here, we indicate that learner experience is more than UX. It certainly includes all aspects of UX, including capturing the quality of a user's experience with a digital technology and examining how easily users perform a task efficiently using a system and how user-friendly, effective, or appealing it is. However, LX also encompasses all aspects related to learning (Jahnke et al., 2020). Based on our work with STP usability heuristics in this research, we see the need to discuss the understanding of LX in the scientific community. From this work here, we suggest the following definition as a useful starting point that includes the technological, pedagogical, and social dimensions.

> *Learning experience (LX) encompasses all aspects of a learner's interaction with: (a) the digital technology/service/space; (b) the pedagogical components, such as course type, learning goals, learning activities, process-based assessment, and learner control; and (c) the social dimension, such as quality of communication forms, collaboration, sociality, social presence, and social interactivity.*

In summary, LX encompasses all aspects of the sociotechnical-pedagogical dimension such as the learner's engagement with the social dimension, the learner's interaction with the digital technology, service, or space, and the learner's interaction with the pedagogical elements.

Limitations

The interrater reliability was lower than some may have expected for well-defined categories, so further research could be done to better define those categories (e.g., train raters). In addition, the raters who assigned the heuristics to problems and rated the severity of the problems were on the same research team. A team from a different academic culture could view some of the problems as falling outside of the 16 final heuristics. Future research is needed.

Also, because the project took place over several semesters, different research team members were involved in different steps of the project. This may or may not impact the results. Future research is needed. Furthermore, only two online courses were evaluated, and both were outside the usual academic credit framework in that they were part of adult learning and an extension division of the university. Further research is needed to test the new heuristics for more traditional courses (e.g., populations of other ages). Further research also may use the new heuristics to score a highly rated course versus a lower rated one, or to compare this set of STP heuristics versus Quality Matters with experienced course evaluators.

Conclusion

In this study, we created a comprehensive set of sociotechnical-pedagogical heuristics (STP heuristics) for evaluating and detecting potential usability issues in online courses. Existing checklists only address specific issues (e.g., system design or pedagogy), while this new set of STP heuristics (Table 4) combines aspects of social elements of online learning, sound pedagogical practices, and technical reliability. The STP heuristics are useful for identifying potential issues in the design or redesign of online courses (Baldwin, Ching & Friesen, 2018). Practitioners and evaluators can use these heuristics as a guide for detecting potential issues and improving the learner experience with online courses. Practitioners (e.g., instructors and instructional designers) can use these heuristics to better plan and organize courses as they build them. Furthermore, these heuristics can be used to identify issues within existing courses as needed. Evaluators (i.e., professionals who assess course quality) can use these heuristics to guide their analysis of technology-heavy courses. While the pedagogical and technological aspects are properly addressed in previous sets of heuristics, the social dimension needs more research. With this sociotechnical-pedagogical set of usability items, we provide a first step that others can use to build upon for further refinement.

Acknowledgements

We are grateful to Jessica Gordon at MU Extension, University of Missouri, for the support for this study. We also thank the doctoral students and research assistants in the Information Experience Lab (ielab.missouri.edu) at the MU School of Information Science and Learning Technologies (SISLT) for supporting the data collection and analysis (in alphabetical order): Courtney Baker, Hao He, Kenneth Kircher, Yen-Mei Lee, Shangman (Eunice) Li, Minh Pham, Matin Qazi, Gayathri Sadanala, Jasmine Tan, and Fan Yu.

Declarations

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

The author(s) received approval from the research ethics review board at the University of Missouri—Columbia, USA for this study.

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Albion, P. (1999a). Heuristic evaluation of educational multimedia: From theory to practice. In J. Winn (Ed.), *Proceedings of the 16th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 9–15). Australasian Society for Computers in Learning in Tertiary Education.

Baldwin, S. J., Ching, Y. H., & Friesen, N. (2018). Online course design and development among college and university instructors: An analysis using grounded theory. *Online Learning*, *22*(2), 157–171. doi:10.24059/olj.v22i2.1212

Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E., Loh, S., & Reeves, T. C. (2002). Usability and instructional design heuristics for e-learning evaluation. In P. Barker & S. Rebelsky (Eds.), *Proceedings from ED-MEDIA 2002: World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 1615–1621). AACE.

Bloom, B. (1956). *Taxonomy of educational objectives handbook I: The cognitive domain*. Addison Wesley Longman, Inc.

Botella, F., Rusu, C., Rusu, V., & Quiñones, D. (2018, Sept.). How novel evaluators perceive their first heuristic evaluation. *Proceedings of XIX International Conference on Human Computer Interaction*. Association for Computing Machinery. https://doi.org/10.1145/3233824.3233835

Boyle, T. (1997). *Design for multimedia learning.* Prentice Hall.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77–101. https://doi.org/10.1191/1478088706qp063oa

Chao, T., Saj, T., & Tessier, F. (2006). Establishing a quality review for online courses. *Educause Quarterly*, *29*(3), 32–39. https://er.educause.edu/articles/2006/1/establishing-a-quality-review-for-online-courses

Chen, B., Chang, Y. H., Ouyang, F., & Zhou, W. (2018). Fostering student engagement in online discussion through social learning analytics. *The Internet and Higher Education*, *37*, 21–30.

Clement, F. J. (1985). Technology & instruction: The rule of seven revisited. *Performance + Instruction*, *24*(2), 6–8. https://doi.org/10.1002/pfi.4150240203

Das, J. (2012). Role of syllabus in higher education: A critical study. *Global Research Methodology Journal*, *7*, 1–2. http://www.grmgrlaranya.com/Journals/7th%20issue/Role%20of%20syllabus%20in%20higher%20education_%20A%20critical%20study.pdf

Demmans Epp, C., Phirangee, K., Hewitt, J., & Perfetti, C. A. (2020). Learning management system and course influences on student actions and learning experiences. *Education Tech Research Dev*, *68*, 3263–3297. https://doi.org/10.1007/s11423-020-09821-1

Dringus, L. P., & Cohen, M. S. (2005, Oct. 19–22). *An adaptable usability heuristic checklist for online courses.* The Frontiers in Education 35th Annual Conference. IEEE.

Douglas, S. (2017, Dec. 6). *7 usability heuristics that all UI designers should know*. UsabilityGeek. https://usabilitygeek.com/usability-heuristics-ui-designers-know/

Fink, S. B. (2012). The many purposes of course syllabi: Which are essential and useful? *Syllabus Journal*, *1*(1), 1–12. https://www.pfw.edu/dotAsset/bdf64113-75c4-487a-ba68-dfcc3044ce9c.pdf

Garrison, D. R., Anderson, T., & Archer, W. (2003). A theory of critical inquiry in online distance learning. In M. Moore & W. Anderson (Eds.), *Handbook of distance education* (pp. 113–128). Lawrence Erlbaum.

Guo, P. J., Kim, J., & Rubin, R. (2014, Mar., 4–5). *How video production affects student engagement: An empirical study of MOOC videos*. The First Association for Computing Machinery Conference on Learning @ Scale Conference. ACM.

Horila, M., Nokelainen, P., Syvänen, A., & Överlund, J. (2002). *Criteria for the pedagogical usability, version 1.0.* Häme Polytechnic and University of Tampere.

Jahnke, I., Schmidt, M., Pham, M., & Singh, K. (2020). Sociotechnical-pedagogical usability for designing and evaluating learner experience in technology-enhanced environments. In M. Schmidt, A. A. Tawfik, I. Jahnke, & Y. Earnshaw (Eds.), *Learner and user experience research: An introduction for the field of learning design & technology.* EdTech Books. https://edtechbooks.org/ux/sociotechnical_pedagogical_usability

Jahnke, I. (2015). *Digital didactical designs: Teaching and learning in CrossActionSpaces*. Routledge.

Jahnke, I., Ritterskamp, C., Herrmann, T. (2005). Sociotechnical roles for sociotechnical systems—A perspective from social and computer sciences. In *Proceedings of the AAAI Fall Symposium on Roles, an Interdisciplinary Perspective: Ontologies, Programming Languages, and Multiagent Systems*. AAAI press.

Kaufmann, R., Sellnow, D. D., & Frisby, B. N. (2016). The development and validation of the online learning climate scale (OLCS). *Communication Education*, *65*(3), 307–321.

Khajouei, R., Gohari, S. H., & Mirzaee, M. (2018). Comparison of two heuristic evaluation methods for evaluating the usability of health information systems. *Journal of Biomedical Informatics, 80,* 37–42. https://doi.org/10.1016/j.jbi.2018.02.016

Lenzner, A., Schnotz, W., & Müller, A. (2013). The role of decorative pictures in learning. *Instructional Science*, *41*(5), 811–831. https://doi.org/10.1007/s11251-012-9256-z

Lim, C. J., & Lee, S. (2007). Pedagogical usability checklist for ESL/EFL e-learning websites. *Journal of Convergence Information Technology, 2*(3), 67–76.

Martin, F., & Bolliger, D. U. (2018). Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. *Online Learning*, *22*(1), 205–222. doi:10.24059/olj.v22i1.1092

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation, 41,* (pp. 85–139). Academic Press.

Moore, J. L., Dickson-Deane, C., & Liu, M. Z. (2014). Designing CMS courses from a pedagogical usability perspective. In A. Benson & A. Whitworth (Eds.), *Research on course management systems in higher education* (pp. 143–169). Information Age Publishing.

Nielsen, J. (1994). *10 heuristics for user interface design*. Nielsen Norman Group. https://www.nngroup.com/articles/ten-usability-heuristics/

Nielsen, J. (2004). *Guidelines for visualizing links*. Nielsen Norman Group. https://www.nngroup.com/articles/guidelines-for-visualizing-links/

Nielsen, J., & Loranger, H. (2006) *Prioritizing web usability.* Pearson Education.

Nokelainen, P. (2006). An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. *Educational Technology & Society*, *9*(2), 178–197. https://www.semanticscholar.org/paper/An-empirical-assessment-of-pedagogical-usability-Nokelainen/ea96b628f440642d72026c14710a67ccd06f41f1#extracted

Obsidian Learning. (2017). *Transforming learning: using video for cognitive, emotional, and social engagement*. https://elearningindustry.com/free-ebooks/transforming-learning-using-video-for-cognitive-emotional-and-social-engagement

Online Learning Consortium. (2018). *OLC OSCQR course design review*. https://onlinelearningconsortium.org/consult/oscqr-course-design-review/

Orcutt, J. M., & Dringus, L. P. (2017). Beyond being there: Practices that establish presence, engage students and influence intellectual curiosity in a structured online learning environment. *Online Learning*, *21*(3), 15–35. doi: 10.24059/olj.v%vi%i.1231

Preece, J. (2001). Sociability and usability in online communities: Determining and measuring success. *Behavior and Information Technology, 20*(5), 347–356. doi: 10.1080/0144929011008468 3

Quality Matters (2018). *Rubrics & Standards*. QM. https://www.qualitymatters.org/qa-resources/rubric-standards

Quinn, C. N. (1996). Pragmatic evaluation: lessons from usability. In A. Christie & B. Vaughan, (Eds.) *Proceedings of Australasian Society for Computers in Learning in Tertiary Education (ASCILITE 96)* (pp. 437–444). University of South Australia.

Quiñones, D., Rusu, C., & Rusu, V. (2018). A methodology to develop usability/user experience heuristics. *Computer Standards & Interfaces, 59*, 109–129. https://doi.org/10.1016/j.csi.2018.03.002

Reeves, T. C. (1994). Evaluating what really matters in computer-based education. In M. Wild (Ed.), *Computer education: New perspectives* (pp. 219–246). Edith Cowan University.

Reeves, T. C., Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E., & Loh, S. (2002, June, 24–29). *Usability and instructional design heuristics for e-Learning evaluation*. ED-Media 2002 World Conference on Educational Multimedia, Hypermedia & Telecommunications.

Robinson, H. A., Sheffield, A., Phillips, A. S., & Moore, M. (2017). Introduction to teaching online: Usability evaluation of interactivity in an online social constructivist course. *TechTrends*, *61*(6), 533–540.

Safie, N. (2007). *An investigation on the relationship between e-learning usability attributes towards motivation to learn*. International Islamic University Malaysia.

Sauro, J. (2014) The relationship between problem frequency and problem severity in usability evaluations. *Journal of Usability Studies, 10*(1)*,* 17–25.

Schade, A. (2014). *Video usability*. Nielsen Norman Group. https://www.nngroup.com/articles/video-usability/

Schuler, D., & Namieka, A. (2017). *Participatory design: Principles and practices*. CRC Press.

Silius, K., & Tervakari, A-M. (2003, May, 8–9). *An evaluation of the usefulness of web-based learning environments. The evaluation tool into the portal of Finnish virtual university*. In V. Peñarrocha (Ed.), International Conference on Network Universities and E-Learning*,* Universidad Politécnica de Valencia.

Sims, R., Dobbs, G., & Hand, T. (2002). Enhancing quality in online learning: Scaffolding planning and design through proactive evaluation. *Distance Education*, *23*(2), 135–148. https://doi.org/10.1080/0158791022000009169

Squires, D., & Preece, J. (1999). Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. *Interacting with Computers*, *11*(5), 467–483.
Stein, J., & Graham, C. (2014). *Essentials of online learning*. Routledge.

Stone, D., Jarrett, C., Woodroffe, M., & Minocha, S. (2005). *User interface design and evaluation*. Morgan Kaufmann.

van der Meij, H., & van der Meij, J. (2013). Eight guidelines for the design of instructional videos for software training. *Technical Communication*, *60*(3), 205–228. https://pdfs.semanticscholar.org/47a8/a0502d3ac8a636277b2ba48cba4e2aa14edd.pdf

van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist*, *38*(1), 5–13. https://doi.org/10.1207/S15326985EP3801_2

Xavier University. (2018). *Online learning heuristics*. https://www.xavier.edu/id/online-blended-classes/

Yousef, A. M. F., Chatti, M. A., Schroeder, U., & Wosnitza, M. (2015). A usability evaluation of a blended MOOC environment: An experimental case study. *International Review of Research in Open and Distributed Learning*, *16*(2), 69–93.

Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker, Jr., J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information and Management*, *43*, 15–27. https://doi.org/10.1016/j.im.2005.01.004

Appendix A

Research Team Members' Involvement in the Study

Step	Step Description	No. Team Members	Team Members
1	Literature review	5	Doc. student 1 Doc. student 2 Doc. student 3 Doc. student 4 Intern student 5
2	Additional items	4	Doc. student 2 Doc. student 4 Doc. student 6 Intern student 7
3	Clustering	2	Doc. student 1 Doc. student 8
4	Selection, adaptation	1	Expert 1
5	Specification, review of the 16 categories	3	Doc. student 6 Doc. student 9 Doc. student 10
6a	Validation (expert review of STP heuristics by frequency)	3	Doc. student 1 Doc. student 8 Doc. student 10
6b	Validation (expert review of STP heuristics by severity of problems)	3	Doc. student 1 Doc. student 8 Doc. student 9
6c	Validation (STP vs. Nielsen and Nokelainen heuristics)	3	Team Nielsen Doc. student 1 Doc. student 6 Doc. student 10
		3	Team Nokelainen Doc. student 1 Doc. student 11 Intern student 12
7	Refinement	2	Doc. student 1 Expert 1

Note. Total research team members were 13. See Figure 1 for more details about steps. Doc. student = doctoral student. Intern students = master students or interns of the lab.