Original citation:
Permanent WRAP url:
http://wrap.warwick.ac.uk/62796

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution- 3.0 Unported (CC BY 3.0) license and may be reused according to the conditions of the license. For more details see http://creativecommons.org/licenses/by/3.0/

A note on versions:
The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: publications@warwick.ac.uk

http://wrap.warwick.ac.uk/
Defining and Evaluating Learner Experience for Social Adaptive E-Learning

Lei Shi

Department of Computer Science, University of Warwick, Coventry, CV4 7AL, UK
lei.shi@dcs.warwick.ac.uk

Abstract
Social adaptive e-learning combines, threads and balances the amount of social and adaptive features for e-learning in order to achieve high-quality Learner eXperience (LX). Evaluating a social adaptive e-learning system is a difficult task due to its complexity. It is crucial to ensure that appropriate evaluation methods and measures are used. A User-centric approach serves the empirical system evaluation using subjective user feedback on satisfaction and productivity as well as the quality of work and support, so as to verify the quality of product, detect problems, and support decisions. This paper proposes a learner-centric evaluation framework, which applies a user-centric approach, aiming to evaluate LX in social adaptive e-learning from the end-user (learner) point of view, taking into consideration both social and adaptive perspectives.

1998 ACM Subject Classification H.5.2 [Information interfaces and presentation]: User Interfaces - Evaluation/methodology

Keywords and phrases Social adaptive e-learning, user-centric evaluation, learner experience

Digital Object Identifier 10.4230/OASIcs.ICCSW.2014.1

1 Introduction
Social adaptive e-learning [31] combines, threads and balances the amount of social and adaptive features for e-learning in order to achieve high-quality overall learner experience. With social features, an e-learning system connects like-minded learners, so they can achieve learning goals via communication and interaction with each other by sharing knowledge, skills, abilities and materials. With adaptive features, an e-learning system delivers learning contents to learners adaptively, namely, the appropriate contents are delivered to the learners in an appropriate way at an appropriate time based on the learners’ needs, knowledge, preferences and other characteristics [32].

Evaluating a social adaptive e-learning system is a difficult task due to its complexity. It is crucial to ensure that appropriate evaluation methods and measures are used. A variety of evaluation methodologies and techniques for adaptive e-learning systems have been proposed. Apart from the traditional prediction accuracy evaluation [15, 20], researchers have reached the agreement on the importance of introducing additional criteria that investigate user experience issues [21] such as trustability [10], enjoyability [13], coverage and serendipity [4], as user satisfaction of a system does not necessarily correlate with the high prediction accuracy. Additionally, since a social dimension has been introduced into adaptive e-learning [3, 29], some evaluation frameworks for social adaptive e-learning have been developed such as [30], which takes into account also social metrics.

Based on existent studies, the study presented in this paper applies a user-centric (learner-centric) approach, aiming to evaluate Learner eXperience (LX) of social adaptive e-learning, from the end-user (learner) point of view, taking into consideration both social and adaptive perspectives. We firstly, in Section 2, review related works including the user-centric
evaluation methodology and existent evaluation frameworks. To propose the learner-centric evaluation framework, we look into general definitions of User eXperience (UX) in Section 3, to simulate the definition of LX and design the evaluation metrics. Section 4 reports a case study using the framework, and finally, Section 5 draws conclusions and future works.

2 Related Work

User-centric evaluation (UCE), as one of the best accepted approaches that identify determinants of User eXperience (UX) issues, serves the empirical system evaluation using subjective user feedback on satisfaction and productivity as well as the quality of work and support [12], aiming at verifying the quality of product, detecting problems and supporting decisions [8]. The nature of UCE makes itself a valuable approach to improve a system and help researchers and engineers implement easier to use and more enjoyable UX, and thus eventually lead to a higher adoption of the system [33].

However, unified frameworks have always been needed to compare the results of different studies. Several recent user-centric evaluation frameworks are proposed to address this issue. Chrysafiadi and Virvou proposed an evaluation framework, called PeRSIVA [7], using Kirkpatrick’s model [16] in combination with a layered evaluation approach to access learners’ satisfaction, performance, progress, behaviour and states. Knijnenburg, et al. proposed a pragmatic procedure that centralised UX including users’ perceptions of a system (perceived system effectiveness and fun), system usage (usage effort and choice difficulty), and outcome of system usage (saturation with the chosen items) [17].

In this paper, based on the definition of Learner eXperience (LX) (Section 3) and existent evaluation frameworks and methods including [2, 7, 16, 17, 23, 34], a new learner-centric evaluation framework for social adaptive e-learning is proposed. This framework aims at evaluating LX from the end-user (learner) point of view taking into consideration both social and adaptive perspectives. It also aims at allowing for comparisons of evaluation results between different social adaptive e-learning systems.

3 Definition and Evaluation of Learner Experience

User eXperience (UX) started to attract designers’ attention since the early 1990s, when it was first coined and brought to wider knowledge by Donald Norman. UX is associated with a wide range of meanings [11], ranging from classical usability to fuzzy and dynamic concepts such as emotional, affective, experiential, hedonic and aesthetic variables [18], which causes itself an elusive notion with many different definitions. Although UX is well discussed in conferences and symposiums, till now there is still no widely share view of its definition. The following are the definitions made by different organisations, researchers and designers just to name a few.

Alben [1] defines UX as all the aspects of how people use an interactive product: the way it feels in their hands, how well they understand how it works, how they feel about it while they are using it, how well it serves their purposes, and how well it fits into the entire context in which they are using it.

Hassenzahl and Tractinsky [14] define UX as a consequence of user’s internal state such as predilections, expectations, needs, motivation, mood, the characteristics of the designed system such as complexity, purpose, usability, functionality, and the context (or the environment) within which the interaction occurs such as social setting, meaningfulness of the activity, voluntariness of use.
ISO 9241-210 [9] defines UX as a person’s perceptions and responses that result from the use and/or anticipated use of a product, system or service, including all the users’ emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviours and accomplishments that occur before, during and after use, influenced by three factors, i.e., system, user and the context of use.

As UX includes a lot of dynamic concepts such as emotional, affective, experiential, hedonic and aesthetic variables, and this is still no widely shared view of its definition, it is even harder to define criteria against which it could be evaluated [24]. The metrics settings for the UX evaluation could be various, depending on the perspectives of both system aims and user needs. For example, for gaming systems, UX metrics may concern the more about the fun and enjoyable experience; for banking systems, UX metrics may concern the more about the secure and trustworthy experience; for online shopping systems, UX metrics may concern the more about customers’ satisfaction on the payment process; and for e-learning systems, UX metrics may concern the more about learners’ engagement of accessing learning materials.

Researchers have been proposing, summarising and classifying the UX metrics (or metrics, criteria, frameworks) in the literature. Van Velsen et al. [33] represented UX metrics such as appreciation, user satisfaction, usability, user performance, intention to use, appropriateness of adaptation, comprehensibility. Morville [22] proposed UX criteria to evaluate if the system is useful, usable, findable, credible, accessible, desirable and valuable. Wu et al.’s UX evaluation framework [34] includes six constructs: flow, perceived technology acceptance, telepresence, performance gains, technology adoption, and exploratory behaviours.

When evaluating UX of more specific types of systems, the metrics are usually tailored to suit the system types’ aims, either by specifying the UX metrics, or introducing some other metrics as the complements.

In the adaptive systems and recommender systems area, several elaborate user-centric evaluation frameworks have been proposed. For example, Pu et al. proposed ResQue [23], which defines a set of metrics that are grouped into four high level layers: perceived system qualities, users’ beliefs, subjective attitudes, and behavioural intentions. For each metric, ResQue also specifies several questions to ask users. Knijnenburg et al.’s framework consists of five dimensions to evaluate including objective system aspects, subjective system aspects, user experience, interaction, and personal and situational characteristics.

In the e-learning area, Ardito et al. [2] proposed four dimensions for the evaluation including presentation, hypermediality, application proactivity, and user activity. And for each of them, both effectiveness and efficiency are considered as the general principles. Liaw and Huang [19] examined the relationships between perceived self-efficacy, perceived anxiety, interactive learning environments, perceived satisfaction, perceived usefulness, and perceived self-regulation, and proposed perceived satisfaction, perceived usefulness, and interactive learning environments as predictors to self-regulation in e-learning environments.

These definitions and metrics indicate that to evaluate social adaptive e-learning systems it is crucial to understand the changes that have occurred due to the merging of techniques into the innovation. Taking into consideration their potential influence on the learning process, it is crucial to characterise learner experience and define the metrics of its evaluation.

In this study Learner eXperience (LX) is defined as a learner’s perceptions and responses resulting from the use and/or anticipated use of an e-learning system - a simulation of UX, whereas more extensive and going beyond the traditional study of skills and cognitive process of users and their behaviours when interacting with an e-learning system. It involves learners’ behaviours, attitudes, believes, sensation, assessment, emotional response obtained
and so on throughout the entire time of using an e-learning system. Based on the discussed
evaluation frameworks and metrics, the following learner-centric scales are developed. Each
of them has several statements in a five-point Likert scale ranging from -2 (strongly disagree)
to 2 (strongly agree), except System Usability Scale (SUS) whose scale ranges from 1 to 5.

**The Learner Belief Scale (LBS)** is designed to test the impact that the learning system
had on the overall LX as perceived by learners. It also evaluates the user-system interaction
considered as helpful in learning as perceived by learners. It consists of ten statements:

- **LBS.01** The system helped me to learn more topics.
- **LBS.02** The system helped me to learn more profoundly.
- **LBS.03** The system helped me to identify my weak points.
- **LBS.04** The system helped me to plan my classwork.
- **LBS.05** The system increased my learning interests.
- **LBS.06** The system increased my learning confidence.
- **LBS.07** The system increased my learning outcome.
- **LBS.08** It was easy to use the system.
- **LBS.09** It was easy to learn how to use the system.
- **LBS.10** It was easy to remember how to use the system.

**The User Interface Scale (UIS)** focuses on the evaluation of learners’ overall feelings
towards a e-learning system’s user interface. The eight UIS statements are:

- **UIS.01** The user interface is familiar to me.
- **UIS.02** The user interface is consistent.
- **UIS.03** The user interface is understandable.
- **UIS.04** The user interface is enjoyable.
- **UIS.05** The user interface is attractive.
- **UIS.06** The user interface is interactive.
- **UIS.07** The user interface is personalised.
- **UIS.08** The user interface is efficient.

**The Content Quality Scale (CQS)** is developed to evaluate learners’ opinion on the
content provided (recommended) by the e-learning system. Its statements include:

- **CQS.01** The content recommended by the system was what I wanted to learn.
- **CQS.02** The content recommended by the system was what I needed to learn.
- **CQS.03** It was easy to recognise the content recommended by the system.
- **CQS.04** I was confident that I would like the content recommended to me.
- **CQS.05** I understood why the system recommended certain content to me.
- **CQS.06** I understood how the content is organised in the system.
- **CQS.07** The content provided by the system was understandable.
- **CQS.08** The content provided by the system was sufficient.
- **CQS.09** The content provided by the system was useful.

**The Socialisation Quality Scale (SQS)** is developed to evaluate students’ opinion on
the e-learning system’s social interaction features. The SQS statements are:

- **SQS.01** It was easy to discuss with the peers.
- **SQS.02** It was easy to share content with peers.
- **SQS.03** It was easy to access the content shared by peers.
- **SQS.04** It was easy to tell peers what I liked/disliked.
- **SQS.05** The statistic numbers (mine and peers’) engaged me to learn more.
- **SQS.06** The system helped me engaged in interacting with peers.
The Behavioural Intention Scale (BIS) is designed to evaluate if the e-learning system could influence learners’ decision to use and reuse the system, as well as to introduce the system to their friends, in order to understand learners’ loyalty. The BIS statements are:

BIS.01 I will use the system again.
BIS.02 I will use the system frequently.
BIS.03 I will tell my friends about the system.

The Perceived Motivation Scale (PMS) is designed based on the Self-Determination Theory (STD) [25], to exam learners’ feelings on their autonomy (question 1-4), competence (question 5-8) and relatedness (question 9-12). The PMS statements are:

PMS.01 I felt in control of my learning process.
PMS.02 I felt interested in using the system.
PMS.03 I felt my learning experience was personalised.
PMS.04 I felt having fun when using the system.
PMS.05 It was easy to understand why I received recommendations.
PMS.06 It was easy to find the content I need.
PMS.07 It was easy to share content with peers.
PMS.08 It was easy to access the shared resources from peers.
PMS.09 It was easy to tell peers what I like/dislike.
PMS.10 It was easy to discuss with peers.

The System Usability Scale (SUS) [6] is a simple, ten-item Likert scale giving a global view of subjective assessments of system usability. It was developed by Brooke in 1996 as a ‘quick and dirty’ questionnaire scale of a given product or service. Its merits make it widely accepted and used in both industry and academic: first, the SUS questionnaire is non-proprietary, making it a cost effective tool; second, the SUS is technology agnostic, and thus flexible enough to evaluate various products and services including hardware, software, websites and applications; third, the SUS questionnaire is relatively quick and easy to use by both experimental subjects and researchers; fourth, the SUS questionnaire provides a single score on a scale, which is easy to understand [37]. The ten items (statements) in the SUS questionnaire are as follows:

SUS.01 I think that I would like to use this system frequently.
SUS.02 I found the system unnecessarily complex.
SUS.03 I thought the system was easy to use.
SUS.04 I thought the system would need support of a technical person to be able to use this system.
SUS.05 I found the various functions in this system were well integrated.
SUS.06 I thought there was too much inconsistency in this system.
SUS.07 I would imagine that most people would learn to use this system very quickly.
SUS.08 I found the system very cumbersome to use.
SUS.09 I felt very confident using the system.
SUS.10 I needed to learn a lot of things before I could get going with this system.

The SUS questions above are all scored on a five-point Likert scale of agreement strength - from strongly disagree to strongly agree. They are alternately positive and negative, so a better system should have higher scores for question 1, 3, 5, 7 and 9, and lower scores for question 2, 4, 6, 8 and 10. The SUS final score is calculated using Equation 1, where $U_i$ presents the score of the $i$-th question. It ranges from 0 to 100, where the higher the final score the better the usability. However, normally a ‘good system’ should have a score between 70 and 80, and an ‘exceptional system’ should have a score greater than 90 [5].
In this study, the SUS questionnaire is used as a 'highly established' measure to evaluate the 'overall usability' of the system. It also contributes to the criterion validity [26] - comparison with other evaluation results using the measures defined in this research.

\[ \text{SUS Score} = 2.5 \times \left( \sum_{n=1}^{5} (U_i - 1) + \sum_{n=1}^{5} (5 - U_i) \right) \]  

4 Case Study

A case study was conducted to demonstrate the usage of the proposed learner-centric evaluation framework. In order to collect data, an experiment was carried out at the Department of Economics, Sarajevo School of Science and Technology, Bosnia and Herzegovina, in December 2013. Twenty students, two observers and one course instructor participated in the one and a half hours online learning session - using Topolor (a social adaptive e-learning system [27]; its user interface is shown as Figure 1) to learn a course on 'Control'. After the initial online session, students were encouraged to further use Topolor to revise the covered materials, for two weeks. After that, students were asked to complete an optional online survey. Out of the twenty students who participated in the online course, fifteen completed the online survey. The results of the survey are shown in Table 1.

The mean values rank between 0.53 and 1.60, with standard deviations ranging from 0.35 to 0.74. All the reported mean values are larger than 0 (the neutral response), suggesting students’ attitudes to be generally positive. Additionally, the results’ Cronbach’s alpha is...
Table 1 Results of LBS, UIS, CQS, BIS and PMS (µ: mean value; σ: standard deviation).

<table>
<thead>
<tr>
<th></th>
<th>µ</th>
<th>σ</th>
<th></th>
<th>µ</th>
<th>σ</th>
<th></th>
<th>µ</th>
<th>σ</th>
<th></th>
<th>µ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBS.01</td>
<td>1.27</td>
<td>0.46</td>
<td>UIS.03</td>
<td>1.00</td>
<td>0.38</td>
<td>CQS.07</td>
<td>1.27</td>
<td>0.46</td>
<td>PMS.01</td>
<td>1.00</td>
<td>0.38</td>
</tr>
<tr>
<td>LBS.02</td>
<td>1.47</td>
<td>0.52</td>
<td>UIS.04</td>
<td>1.07</td>
<td>0.46</td>
<td>CQS.08</td>
<td>1.27</td>
<td>0.46</td>
<td>PMS.02</td>
<td>1.27</td>
<td>0.46</td>
</tr>
<tr>
<td>LBS.03</td>
<td>0.93</td>
<td>0.46</td>
<td>UIS.05</td>
<td>0.87</td>
<td>0.74</td>
<td>CQS.09</td>
<td>1.33</td>
<td>0.49</td>
<td>PMS.03</td>
<td>1.40</td>
<td>0.51</td>
</tr>
<tr>
<td>LBS.04</td>
<td>1.00</td>
<td>0.65</td>
<td>UIS.06</td>
<td>0.80</td>
<td>0.56</td>
<td>SQS.01</td>
<td>0.93</td>
<td>0.46</td>
<td>PMS.04</td>
<td>1.40</td>
<td>0.51</td>
</tr>
<tr>
<td>LBS.05</td>
<td>0.87</td>
<td>0.74</td>
<td>UIS.07</td>
<td>0.93</td>
<td>0.46</td>
<td>SQS.02</td>
<td>1.13</td>
<td>0.35</td>
<td>PMS.05</td>
<td>1.20</td>
<td>0.41</td>
</tr>
<tr>
<td>LBS.06</td>
<td>1.07</td>
<td>0.70</td>
<td>UIS.08</td>
<td>0.80</td>
<td>0.41</td>
<td>SQS.03</td>
<td>0.93</td>
<td>0.46</td>
<td>PMS.06</td>
<td>1.40</td>
<td>0.51</td>
</tr>
<tr>
<td>LBS.07</td>
<td>0.87</td>
<td>0.52</td>
<td>CQS.01</td>
<td>1.13</td>
<td>0.52</td>
<td>SQS.04</td>
<td>1.12</td>
<td>0.41</td>
<td>PMS.07</td>
<td>1.33</td>
<td>0.49</td>
</tr>
<tr>
<td>LBS.08</td>
<td>1.00</td>
<td>0.53</td>
<td>CQS.02</td>
<td>0.73</td>
<td>0.59</td>
<td>SQS.05</td>
<td>1.13</td>
<td>0.35</td>
<td>PMS.08</td>
<td>1.27</td>
<td>0.59</td>
</tr>
<tr>
<td>LBS.09</td>
<td>0.87</td>
<td>0.35</td>
<td>CQS.03</td>
<td>0.93</td>
<td>0.46</td>
<td>SQS.06</td>
<td>1.60</td>
<td>0.51</td>
<td>PMS.09</td>
<td>1.80</td>
<td>0.41</td>
</tr>
<tr>
<td>LBS.10</td>
<td>0.87</td>
<td>0.52</td>
<td>CQS.04</td>
<td>1.60</td>
<td>0.51</td>
<td>BIS.01</td>
<td>1.20</td>
<td>0.41</td>
<td>PMS.10</td>
<td>0.53</td>
<td>0.52</td>
</tr>
<tr>
<td>UIS.01</td>
<td>0.87</td>
<td>0.52</td>
<td>CQS.05</td>
<td>1.27</td>
<td>0.46</td>
<td>BIS.02</td>
<td>1.07</td>
<td>0.46</td>
<td>PMS.11</td>
<td>0.93</td>
<td>0.59</td>
</tr>
<tr>
<td>UIS.02</td>
<td>0.87</td>
<td>0.52</td>
<td>CQS.06</td>
<td>1.47</td>
<td>0.52</td>
<td>BIS.03</td>
<td>1.60</td>
<td>0.51</td>
<td>PMS.12</td>
<td>1.20</td>
<td>0.41</td>
</tr>
</tbody>
</table>

0.864 (>0.8), suggesting a high level of result reliability. The SUS score is 73.67 out of 100 (σ=9.01, median=72.50), indicating Topolor is a ‘good system’, according to Section 3.7.

5 Conclusion and Future Work

This study has proposed a new learner-centric framework for evaluating Learner eXperience (LX) in social adaptive e-learning. The new recommended definition of LX - learner’s perceptions and responses resulting from the use and/or anticipated use of an e-learning system, is a simulation of User eXperience (UX) from Human-Computer Interaction (HCI) research area, which is more extensive and going beyond the traditional study of skills and cognitive process of users and their behaviours when interacting with an e-learning system. Seven learner-centric scales have been defined including LBS (Learner Belief Scale), UIS (User Interface Scale, introduced), CQS (Content Quality Scale), SQS (Socialisation Quality Scale), BIS (Behavioural Intention Scale), PMS (Perceived Motivation Scale) and SUS (System Usability Scale). They involve learners’ behaviours, attitudes, believes, sensation, assessment, emotional response obtained and so on throughout the entire time of using an e-learning system.

Following the definition of LX and its metrics, a case study was presented to demonstrate the use of the proposed learner-centric evaluation framework. The evaluation results illustrated generally positive students’ attitudes of using Topolor, a social adaptive e-learning system. The main limitation of the case study is the low number of participants, although Cronbach’s Alpha suggests a high level of reliability of the results. However, Topolor has been opened to public (www.topolor.com), with larger student cohorts expected in the near future, allowing for feedback, use data and suggestions collecting, in further studies.

Whilst subjective and objective measures could be substitutes for each other, both empirical and analytical studies have suggested that the two types of measurement are equally important and thus both of them should be considered. Therefore, in the follow-up work, objective measures will be integrated into the proposed evaluation framework. Learners’ usage data will be collected using data-logging techniques and analysed using various data mining and visualisation tools. In fact, this work has been already initialised, detailed in [28].
References