Reflections on the Evaluation of Adaptive Learning Technologies

Dietrich Albert\textsuperscript{1,2} & Christina M. Steiner\textsuperscript{1}

\textsuperscript{1}Knowledge Management Institute, Graz University of Technology
\textsuperscript{2}Department of Psychology, University of Graz
Graz, Austria

e-mail: {dietrich.albert, christina.steiner}@tugraz.at

Abstract—The key characteristic of adaptive learning technologies implies specific challenges for evaluation. Despite an increasing awareness of the need for sound evaluations, the current evaluation practice still suggests considerable weaknesses in the approaches applied and calls for further research on appropriate and sophisticated evaluation methodologies. This paper aims at further advancing the maturity of evaluation approaches for adaptive (learning) technologies by stimulating discussion and dialogue on future research directions.

Keywords—adaptive learning technology; adaptation; evaluation

I. THE CHALLENGES

Adaptive learning technologies acknowledge the fact that learners may feature different goals, starting conditions, knowledge, experience, interests etc. and try to respond to this diversity by tailoring learning contents, feedback and user interface to the needs of the individual user\footnote{1}[2]. To demonstrate the potential and benefit of adaptive learning technologies sound evaluation is needed. The main challenge of evaluating an adaptive learning technology lies in its key characteristic, namely the adaptation to the learner. It results in individual experiences and interactions with the system for each individual user - and exactly this personalised output complicates evaluation\footnote{3}.

When evaluating an adaptive learning technology, relevant evaluation questions are: does adaptivity improve the system, do users prefer the adaptive version and features, does adaptation improve interaction between user and system etc.\footnote{4}. A crucial question, of course, is also whether adaptation leads to a benefit for learning.

II. THE STATUS QUO

In recent years, the importance of evaluating adaptive systems has been increasingly acknowledged and gradually more evaluation studies are reported\footnote{5}.

Traditional evaluation practice on adaptive systems makes use of comparative approaches realising ‘with and without designs’. This means that an adaptive instance of a system is compared to a non-adaptive counterpart\footnote{6}. As, though, adaptation is naturally designed into a system the non-adaptive version will not be optimal in any case. Furthermore, no insights are provided into why the adaptive system is better (or not). Another option is to compare the adaptive system with a different learning technology or with traditional learning methods. This approach, however, confounds the specific effects of adaptation with the differences in look and feel and/or in contextual conditions, not to mention the possible novelty effect. Comparative evaluation can be considered a useful approach when an existing non-adaptive learning system or digital library for research and education is complemented by adaptive functionality (as e.g. the case in the CULTURA project\footnote{8}). A recent approach of comparative evaluation adopted a yoked-control design to assure a scientifically sound comparison\footnote{7}.

Another evaluation approach is to consider adaptivity as an integral part of the system\footnote{9}, without any dependence on the presence and comparison with a non-adaptive condition. Such criteria-based evaluation tries to identify basic criteria for giving evidence whether adaptation works and provides an added value. The problem with this approach, however, is that it is not possible to assess the degree to which different factors contribute to success or failure.

A lot of research in recent years has been dedicated to elaborating approaches of layered evaluation, where the adaptation process is broken down into its individual components for separate evaluation, rather than treating it as a singular process. A range of different frameworks has been devised, which differentiate adaptation and evaluation layers at diverse levels of granularity. The layered evaluation approach suggested by\footnote{10} distinguishes two main phases of adaptation, which should be investigated separately in evaluation: interaction assessment and adaptation decision making. The most recent framework unifies previous approaches and comprises a more fine-grained decomposition of adaptation and according evaluation\footnote{5}.

Despite an increasing awareness of the importance of evaluation, current frameworks for evaluating adaptive learning systems and user-adaptive systems, in general, can be considered not yet fully elaborated\footnote{3,5}.

III. CURRENT TRENDS AND FUTURE DIRECTIONS

Three perspectives are crucial for a further evolution of evaluation methods and practice in the context of adaptive learning technologies: process-oriented evaluation, theory-driven evaluation, and continuous evaluation.

A. Process-oriented Evaluation

Process-orientation refers to the consideration of evaluation as an inherent part of the development lifecycle of an adaptive system. Evaluation needs to be done iteratively, accompanying the design and development of an adaptive system from the beginning till the end. This covers the elicitation of users’ requirements (requirements analysis), the identification of problems and issues in initial implementations, to be fed back into the
development cycle (formative evaluation), and the investigation of the overall quality and added value of a technology (summative evaluation). The current focus of adaptive systems’ evaluations is clearly on summative evaluation [5]. Increased effort needs be put on the formative evaluation of early prototypes and implementations. In particular, the early involvement of users as co-researchers and co-designers needs to be established to ensure that users’ needs are sufficiently reflected and to stimulate motivation and acceptance to use the adaptive technology [11][12]. This early and continuous dialogue with users is at the core of the evaluation approaches in the European projects GRAPPLE [13] and the CULTURA [8].

Another perspective on process-oriented evaluation is to consider the whole process of creating and delivering adaptation. This involves the evaluation of authoring tools for the creation of adaptive courses [14] in addition to the analysis of the adaptive learning experience itself.

B. Theory-driven Evaluation

For evaluating adaptive learning systems a theory-driven approach to evaluation is critical for receiving sound and meaningful outcomes. Theory-driven may thereby be considered as a systematic approach towards evaluation in terms of differentiating between different, theoretically inspired dimensions to be considered in evaluation, e.g. development phases [3], adaptation layers [10], expected outcomes for learners etc. These dimensions may be used to set up a multidimensional model that can be applied as a basis for systematic evaluation studies. Psycho-pedagogical theories are used as a basis for realizing learning and educational adaptation in an adaptive learning system. Evaluation needs to be oriented alongside these psycho-pedagogical theories in order to investigate whether the theoretical knowledge has been appropriately incorporated. The scientific soundness and value of the theoretical structures, assumptions, and processes underlying the adaptive learning technology need to be evaluated. This may consist in investigating whether the knowledge and competence structures [15] assumed for a learning domain adequately reflect the real world, whether the applied natural language processing methods lead to valid interpretations of text, whether a didactic strategy has been appropriately translated into a learning technology or which of different forms of assistance or interaction are most appropriate [16]. Examining the scientific soundness of the theoretical approaches feeding into adaptivity is also part of the evaluation methodology in the CULTURA project [8]. Only if an adaptive system grounds on valid methods the resulting adaptation will be of value – otherwise, it might be even worse than a ‘one size fits all’ approach.

C. Continuous Evaluation

Finally, we see great potential for (r)evolving the future of evaluating adaptive learning technologies and adaptive systems, in general, in continuous approaches to evaluation. Formal theoretical approaches (e.g. [15]) may be adapted for continuous and semi-automatic evaluation of adaptation. This may involve the use of non-invasive assessment techniques [17] and the exploitation of user interaction and log-data. In addition, possible ways of gathering explicit on-line feedback from users while dealing with the system need further research. This might consist in the application of mood barometers or the utilisation of tagging functionalities. The analysis of learner traces based on implicit and explicit data collection can help to establish a more comprehensive picture of the process of adaptation and its effects. In CULTURA an evaluation tool for continuous and longitudinal analyses of adaptation quality will be researched and developed [8].

ACKNOWLEDGMENT

The work presented in this paper is partially supported by the European Commission under the Seventh Framework Programme in the ICT research priority (GRAPPLE, project no. 215434 and CULTURA, project no. 269973).

REFERENCES