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.

An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students

Petri Nokelainen
Research Centre for Vocational Education, University of Tampere, FIN-13101 Hämeenlinna, Finland
http://www.uta.fi/u/petri.nokelainen
Tel: +358 40 557 4994
Fax: +358 3 6145 6111
petri.nokelainen@uta.fi

ABSTRACT

This paper presents the pedagogical usability criteria for evaluating the digital learning material. Pedagogical aspects of designing or using digital learning material are much less frequently studied than technical ones. Further, there are relatively few inventories measuring subjective end-user satisfaction with the pedagogical aspects of digital learning material and not a single inventory has undergone a rigorous process of empirical psychometric testing. They include the following components: 1. Learner control, 2. Learner activity, 3. Cooperative/Collaborative learning, 4. Goal orientation, 5. Applicability, 6. Added value, 7. Motivation, 8. Valuation of previous knowledge, 9. Flexibility and 10. Feedback. The pedagogical usability criteria have been operationalized into an on-line Likert -scale self-rating Pedagogically Meaningful Learning Questionnaire (PMLQ) that has 56 items. In the PMLQ, separate items have been developed to measure the usability of the learning management system (LMS) and the learning material (LM). When evaluating the usability of a LMS, it is possible in the pedagogical context to evaluate the kind of learning material it enables the users to produce. Evaluation of the usability of a LM is based on a fact that the learning content is based on a certain learning goal or goals. The criteria of pedagogical usability presented here have undergone two-step psychometric testing process using empirical samples of 5th and 6th grade elementary school students (n = 66 and n = 74). Students evaluated one LMS and four LM's with PMLQ. Results supported the existence of theoretical dimensions of the criteria. The PMLQ was able to capture differences in the pedagogical usability profiles of the learning modules. Generalizability of the pedagogical usability criteria to other domains is limited by the small sample size, the small age range of respondents' and the small number of learning material evaluated. However, empirical studies that aim at evaluating a more representative set of learning material in different domains targeted for both adolescent and adult learners are currently conducted.

Keywords

Pedagogical issues, Usability, Evaluation methodologies, Evaluation of CAL systems, Authoring tools and methods

Introduction

New software has eliminated some previously prerequisites student skills, while simultaneously creating new requirements for the computer systems they are using. The usability analysis of learning material is intended to influence the process by which the correct applications and target groups find one another so that a student can focus his/her energy on the content of the learning material rather than technical issues caused by the software design or user interface.

Several sets of recommendations for the evaluation of technical usability have been developed over the last twenty years (e.g., Shneiderman, 1998; Chin, Diehl & Norman, 1988; Nielsen, 1993; 1994; Lin, Choong & Salvendy, 1997; Preece, Rogers & Sharp, 2002; Chalmers, 2003; Tognazzini, 2003). However, pedagogical aspects of designing or using digital learning material are much less frequently studied than technical ones. This paper presents criteria for evaluating the pedagogical usability of digital learning material. The purpose of the criteria is not to brand any learning material as "good" or "bad," but to help users choose the most suitable alternative for any particular learning situation.

When evaluating *technical usability*, the basic assumption is that it should be easy to learn to use the central functions of the system and the functions are efficient and convenient in use. Another assumption is that error responses to incorrect operation of the software should help teach the user to use the system as intended so that the error will not be repeated. When evaluating *pedagogical usability*, the assumption is that the designers of the learning platform or learning unit were guided by either a conscious or subconscious idea of how the functions of the system facilitate the learning of the material it is delivering. Examples of the range of learning theories that influence design paradigms are objectivism (instructivism, behaviourism), and constructivism (focus on learner, learner's active role in learning and learning from experience).

The criteria presented here have been operationalized into an on-line application (Nokelainen, 2004a) that applies a Likert –scale self-rating questionnaire with 56 items (see Appendix). In the Pedagogically Meaningful Learning Questionnaire (PMLQ), separate items have been developed to measure the usability of the learning platform and the learning unit it is delivering. When evaluating the usability of a specific *learning platform* (learning management system), we can evaluate how easy it is to use the platform itself (technical usability), or what kind of learning material it enables the users to produce (pedagogical usability). When evaluating the usability of a *learning unit* or a *learning object*, we assume that each learning unit has its own interface relating to the content, and learning content based on a certain learning goal. When evaluating the pedagogical usability of a learning unit, we must try to control the effect of the pedagogical solutions of a learning platform.

The compilation of criteria for usability of digital learning material presented here is based on the following preparatory analyses: 1. Charting of systems and taxonomies developed for the evaluation of digital learning materials, 2001-2002, 2. Analysis of learning theories suitable as the basis of evaluation, 2001-2002, 3. Compilation of the first draft of criteria for pedagogical usability, 2001-2002, and 4. Testing of the first draft of criteria for pedagogical usability with empirical samples, 2001-2002.

Central Concepts

Digital learning material means in this research all material that is designed for educational purposes, published in a digital form and intended to be accessed by computer.

A Learning object (LO) is, as the name implies, the smallest reasonable unit of learning material. Examples of this are a pronunciation sample from an English language-teaching program or an animation clip that describes how a dangerous procedure is completed safely.

Learning material (LM) can include such things as 1) a WWW page introducing the effects of image compression on the quality of an image, 2) a JAVA applet to practice touch-typing, 3)an individual learning module in a learning platform, the topic of which might be, for example, positioning decimal numbers on a continuum. Regardless of the type, each individual learning material has its own user interface, the usability of which can be evaluated, as well as a definable learning goal.

A Unit of learning material (ULM) consists of several individual learning materials connected by a common main goal. A ULM is a concept that combines several learning objects (and possible meta, history and relation data) related to the same subject as a single entity.

A Virtual learning environment (VLE) is an application that uses learning materials, or units of learning material. There are numerous commercial and non-commercial virtual learning environments available. They can be divided into two groups according to whether they include prepared learning material (e.g., eWSOY's OPIT) or offer a presentation service for learning material provided by the user (e.g., BlackBoard, Belvedere, Dyn3W, Future Learning Environment, Knowledge Forum, WebCT)

Usability can, according to Keinonen (1998, 62) be defined as a characteristic related to 1)the product's design process, 2) the product itself, 3) use of the product, 4) user experiences of the product or 5) user expectations. In the present research, usability is understood primarily as the product's usability attributes, which are measured through subjective user experiences with a self-evaluation questionnaire.

The operationalisation of the usability attributes in this study is done according to the model presented by Nielsen (1990, 148). Figure 1 shows that the top-level concept is system acceptability. Acceptability is further divided into two parts: 1) Social acceptability and 2) Practical acceptability. The implementation of a system that serves as a presentation platform reveals the views of its developers about both social and practical acceptability. The example given by Nielsen to illustrate social acceptability (*ibid.* 147-148) is a curriculum that lets only the teacher make changes that increase viewpoints on the studied subject, which could in some societies be seen as inappropriately emphasising the teacher's authority and restricting the students' individual knowledge formation. The tone of Nielsen's later example of social acceptability (1993, 24) puts more emphasis on the ethical and moral choices of the program developers: Is it right that people applying for unemployment benefits would have to use an application form that attempts to use their input (questions) to reveal those who are applying for benefits for invalid reasons? According to Nielsen (1990, 148), practical acceptability means, among other things, cost, compatibility, reliability and usefulness of the system. Nielsen states (1990, 148) that usefulness can be further divided into utility and usability. Utility refers to the ability of the system to generally provide functions that correspond with the needs of the users and usability refers to how well the users are able to use the functions offered by the system (*ibid.* 148-149). According to Nielsen's examples (1990, 151), an entertainment application (e.g., a game) is utilisable for users if they enjoy themselves with it, and learning material is utilisable if students learn with it. Nielsen (*ibid.* 151) further divides usability into learnability, efficiency, memorability, errors, and satisfaction, calling them *usability attributes*. *Learnability* depends on how long beginners' use a system before they learn the essential skills necessary to perform their tasks. *Efficiency* refers to how well experienced users can operate an application after they have mastered it. *Memorability* means the ability of an occasional user who has previously used the system to remember its operational principles. *Errors* are divided into two groups: less serious errors that disturb the work of the user and serious errors that, for example, endanger the preservability of the users' outputs. *Satisfaction* is a subjective judgment by the user.

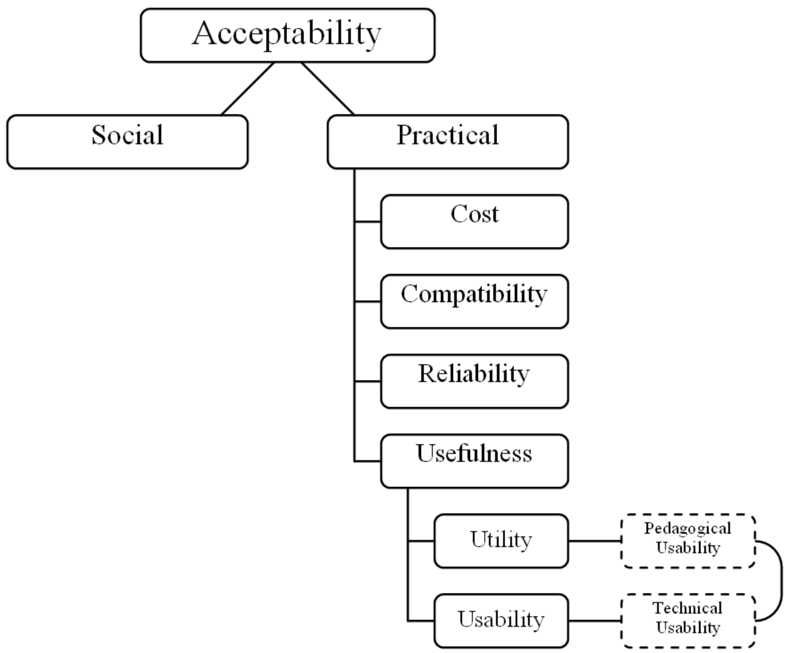

*Figure 1***.** Conceptual mapping of the technical and pedagogical usability to the Nielsen's usability model (Adapted from Nielsen 1990, 148)

Pedagogical usability, which is defined in this research according to Nielsen's classification (1990, 148), is a sub-concept of utility, and *technical usability* is a sub-concept of usability (Figure 1). Thus, in addition to the dialogue between a user and a system, the pedagogical usability of a system and/or learning material is also dependent on the goals set for a learning situation by the student and teacher.

Evaluation in this research is a subjective judgment made by the user of digital learning material. The evaluation criteria are applied using a self-evaluation questionnaire that employs a Likert scale 1 (strongly disagree) – 5 (strongly agree), 6 (not applicable), 7 (don't know). The teacher and learners/students each have their own versions of the evaluative statements.

Pedagogical Usability

Learning is a largely unobservable and uncontrollable process that happens all the time. Attributes attached to learning such as 'effectiveness' or 'informality' are both difficult to understand and measure. The term education, on the other hand, is easier to analyse because it is more often bound to observable artefacts such as books or computer programs that are related to a curriculum. Some researchers term education that is tied to a set curriculum as ‘formal,’ and education that is more conversation based in nature as ‘informal’ (Livingstone, 2000). Although we will go no further with this topic, one may ask what portion of measurement error of the pedagogical aspects of usability is related to the difference between the formal and informal aspects of education.

Education can be provided in groups or according to individual needs as teacher directed, cooperative, or individual practice. When computers and digital learning material are used in a learning situation, it is expected that this is done to introduce identifiable added value to the learning in comparison to, for example, printed material, and material produced by the teacher or the students themselves. The expectation of “added value,” not merely “reaching the same level,” is demonstrated by the fact that each group of students is under the charge of an educational expert who has primary responsibility for the provision of education that agrees with the general goals, and further supported by computers and software that require not only capital investment but also continuous maintenance. In the case of distance education, the equation is different and so it may be that reaching the same level as in contact teaching is considered acceptable. In this context, the level of educational quality desired is at least that of the average performance in a good quality contact teaching context – it is unlikely that one would want to match digital material with poor learning materials and teaching practices! In the following the problematic concepts of the quality of education and learning are discussed in more detail.

Many of the “innovations” that have been executed in computer environments have their equivalents in the world of traditional education. For example, instead of a “mind map” type of computer application that collects arguments and counter-arguments, cooperative knowledge construction can be achieved with large pieces of paper on the walls of a classroom, on which various points and opinions that have been raised during a week can be recorded. The added value of computers in this case is technical, because it gives the students a chance to work simultaneously on several different things (for example in different subjects) and save each phase of their work on their own workspace. In a classroom, the activity is more limited and inflexible because the space on the walls of the classroom is limited. Then again, if the unfinished work and problems were visible at all times, the students could work on them in the background for the whole time that they spend in the classroom. A counter-argument to this is the Sartrean thought about comprehensive thinking that follows the learner from school to home. In this case, a computer-based application allows the learners to continue their work at home, in a library or in a coffee shop with the help of a computer network.

Developing the New Criteria for Pedagogical Usability

We analysed the criteria that have been developed for the factors in learning materials that promote learning (Reeves, 1994; Squires & Preece, 1996; Quinn, 1996; Albion, 1999; Squires & Preece, 1999; Horila, Nokelainen, Syvänen & Överlund, 2002). The summary of different criteria is presented in Table 1. The small number of existing models shows that pedagogical aspects of designing or using digital learning material are much less frequently studied than technical ones. The table shows that one criteria is a theoretical model (Squires & Preece, 1996), four are theoretical models with heuristic checklists (Reeves, 1994; Quinn, 1996; Albion, 1999; Squires & Preece, 1999) and only one is inventory measuring subjective end-user satisfaction with the pedagogical aspects of digital learning material (Horila et al., 2002). Further, not a single heuristic checklist or inventory had undergone a process of empirical psychometric testing.

The most common pedagogical usability features of the criteria presented in Table 1 are “learner control”, “possibility for cooperative or collaborative learning activities”, “explicit learning goals”, “authenticity of learning material” and “learner support (scaffolding)”.

In this study, we created new criteria for the assessment of the pedagogical usability of digital learning materials, as the earlier research work had not addressed all the relevant issues of the topic neither on a theoretical nor practical level. The existing criteria neglect partially the role of learner’s activity, added value of digital learning material, learning motivation and feedback related to user input. We found that none of the existing criteria included such concepts as valuation of previous knowledge and role of pre-testing and diagnostics. Our model includes ten dimensions as follows: 1. Learner control, 2. Learner activity, 3. Cooperative/Collaborative learning, 4. Goal orientation, 5. Applicability, 6. Added value, 7. Motivation, 8. Valuation of previous knowledge, 9. Flexibility and 10. Feedback.

Next we proceed to describe the theoretical structure of each dimension. The model presented here contains 51 sub dimensions. To get the picture of the overall structure of the criteria, the reader is encouraged to see the pedagogical usability sub dimensions in the Appendix 2.

1 Learner Control

When learning a new topic, the learner's memory should be burdened to an optimal level (Miller, 1956; Shneiderman, 1998, 355). It is difficult to define a universally optimal level (generally people can have 5 to 9 items in their short term work memory), but it is certainly helpful to break down the material to be learned into meaningful units (Wilson & Myers, 2000). In so-called structured learning materials, the learned material has already been broken down into meaningful units from the point of view of the students (albeit usually by the teacher). Such a "one-size-fits-all" approach has been criticized for the fact that the students are required to adjust their learning to fit the teacher's conception of the best way of learning certain material (Jonassen, Myers & McKillop, 1996). The factor structure of this dimension is presented in Table 2.

2 Learner Activity

A teacher's "didactic role" in a learning situation may strongly scaffold the learners' own activity, and, correspondingly, the learners' independent activity may be increased when the teacher stays in the background, as a "facilitator" (Reeves, 1994). Learners' activity is determined in large measure but the characteristics of the learners themselves, but the learning material can affect it through assignments that support student activity by being interesting and based on real life. An alternative approach to structured material is problem-based learning (e.g., Savery & Duffy, 1995). In this case, the teacher gives the students a certain amount of source material from which the students (individually or in groups) construct their own conception of the topic to be learned. However, even in this approach all available information (i.e. "source material") has usually already been structured by someone else! Problem-based learning or accessory software also increases student activity (Jonassen, Peck & Wilson, 1999). Items 8, 9 and 10 reflect more collaborative nature of learning (Forman & Cazden, 1985; Palloff & Pratt, 2005; Tudge & Rogoff, 1989). Study of collaborative activity in computer-based settings, often cited as computer supported collaborative learning (CSCL), can enhance our understanding of the processes of productive interaction (e.g., Littleton & Häkkinen, 1999). The factor structure of this dimension is presented in Table 3.

3 Cooperative/Collaborative Learning

Cooperative and collaborative learning means studying with other learners to reach a common learning goal. According to Barab and Duffy (2000, 27-28), learners are moving away from acquisition metaphor (i.e., acquiring knowledge that is constituted of symbolic mental representations) to participation metaphor (i.e., knowledge is considered fundamentally situated in practice). Instead of acquiring personal knowledge, learners construct knowledge as members of communities in practice (Lave & Wenger, 1991). To be more specific, cooperative learning is more structured than collaborative learning, as the teacher has the control. Learning takes place in groups in which the members gather and structure information, in which case the system or learning material should offer the learner tools that can be used to communicate and negotiate different approaches to a learning problem (Jonassen, 1995). Through the use of computer-supported learning material it is possible to practice cooperative learning so that all students are connected to each other over a distance, for example through discussion groups and chat forums (Quinn, 1996; Reeves 1994). In a system or learning material supporting cooperative knowledge construction (KB, knowledge building or KC, knowledge construction) may include, for example, a visual tool that the learners can use to fashion simultaneous mind maps of the topic. Often such a system also includes tools for social navigation (Mayes & Fowler, 1999; Kurhila, Miettinen, Nokelainen & Tirri, 2002) with which a learner gains information about what the other learners have done (asynchronic social navigation) or are presently doing (synchronic social navigation).

4 Goal Orientation

As learning is a goal-oriented activity, goals and objectives should be clear to the learner (Quinn, 1996). The best results are attained when the goals of the learning material, teacher and student are closely aligned. The goals may vary from concrete (for example the basic techniques of first aid) to abstract (learning material aimed to develop the appreciation of modern arts) (Reeves, 1994). If the learners themselves do not set the goals, their meaningfulness should be justified from the point of view of motivation. The students should have a chance to pursue their own interests in relation to the learning goals. According to instructivist learning theory, learners should be introduced to only with a few, clearly specified goals at any one time. According to constructivist learning theory, the goals should be clearly defined, but they have to originate, as much as possible, with the learners themselves. (Wilson & Myers, 2000.) If the goals do not originate with the students, their meaning should be explained to them. The students should have a chance to make choices with respect to the course of their studies in relation to the learning goals. For example, in some learning materials, the assignments are graded and granted points and thus the student can access the next assignment after garnering enough points from the previous assignments. The downside of the point system is the fact that the students are less able to complete the material in the order of their own preference, or bypass areas that they already know.

*Table 1***.** Summary of the pedagogical usability criteria research

Reeves (1994) "Pedagogical dimensions" [b]	**Quinn (1996)** "Educational design heuristics" [b]	**Squires & Preece (1996)** "JIGSAW model" [a]	**Albion (1999)** "Content heuristics" [b]	**Squires & Preece (1999)** "Learning with software heuristics" [b]	**Horila, Nokelainen, Syvänen & Överlund (2002)** "Pedagogical usability of digital learning environments" [c]
1. Learner control	1. Clear goals and objectives	1. Specific learning tasks	1. Establishment of context	1. Appropriate levels of learner control	1. Learnability
2. Pedagogical philosophy	2. Context meaningful to domain and learner	2. General learning tasks	2. Relevance to professional practice	2. Navigational fidelity	2. Graphics and layout
3. Underlying psychology	3. Content clearly and multiply represented and multiply navigable	3. Application operation tasks	3. Representation of professional responses to issues	3. Match between designer and learner models	3. Technical requirements
4. Goal orientation	4. Activities scaffolded	4. General system operation tasks	4. Relevance of reference materials	4. Prevention of peripheral cognitive errors	4. Intuitive efficiency
5. Experiential value (Authenticity)	5. Elicit learner understandings		5. Presentation of video resources	5. Understandable and meaningful symbolic representation	5. Suitability for different learners and different situations
6. Teacher role	6. Formative evaluation		6. Assistance is supportive rather than prescriptive	6. Support personally significant approaches to learning	6. Ease of use: Technical and pedagogical approach
7. Program flexibility	7. Performance should be 'criteria-referenced'		7. Materials are engaging	7. Strategies for the cognitive error recognition, diagnosis and recovery	7. Interactivity
8. Value of errors	8. Support for transference and acquiring 'self-learning' skills		8. Presentation of resources	8. Match with curriculum	8. Objectiveness
9. Cooperative learning	9. Support for collaborative learning		9. Overall effectiveness of materials		9. Sociality
10. Motivation					10. Motivation
11. Epistemology					11. Added value for teaching
12. User activity					
13. Accommodation of individual differences (Scaffolding)					
14. Cultural sensitivity					

[a] = Theoretical model. [b] = Theoretical model and heuristic checklist. [c] = Theoretical model and subjective end-user inventory.

Table 2. Factor structure of the first pedagogical usability dimension (1. Learner control)

	Target group					
	Child		Teacher		Adolescent/adult	
Factors	Items LMS	Items LM	Items LMS	Items LM	Items LMS	Items LM
1.1 Minimum memory load		47		47		47
1.2 Meaningful encoding		48, 49		48, 49		48, 49
1.3 Take responsibility for one's own learning		1		1		1
1.4 User control	2	2	2	2	2	2
8.2 Valuation of Previous Knowledge -> Elaboration		*40*		*40*		*40*

Note. LMS = Learning management system. LM = Learning module/material. See Appendix for the item labels. Italicised items show relationship to the other usability dimensions.

5 Applicability

The approach taken in learning material should correspond to the skills that the learner will later need in everyday and working life (Jonassen, Peck & Wilson, 1999; Quinn, 1996). The skills or learned knowledge should be transferable to other contexts (Quinn, 1996; Reeves, 1994). Learning something new is most effectively accomplished through so-called learning by doing methods that involve practical tasks. Learning by doing has been found to be an effective learning method for both practical (changing some part of a paper pulp machinery) and in abstract (integration in mathematics) issues. Learning material should always be at an appropriate level from the point of view of a learner's learning process (Wilson & Myers, 2000). For example, students in the early grades of comprehensive school have limited ability to adopt abstract concepts, but it becomes far easier for fifth and sixth graders. It is possible to adapt learning material to meet the needs of a student for, example, by periodically asking for feedback about the experienced difficulty of the material. In a dynamic system, it is possible to follow a student's activity, such as the time it takes to finish assignments and the number of mistakes, and adapt future tasks according to this information. Learning material should be planned and executed in cooperation with both groups of end-users (students and teachers); systems produced by experts rarely meet their everyday needs without modifications. In the planning of learning material special attention can be given to those issues that are most likely to cause problems for the learners (for example the adoption of a new concept) and build support structures that the students can use if needed. The system can also observe the learner's activity and proactively offer hints (prompting and fading, Hannafin & Peck, 1988, 47). When students are in a so-called ZPD (zone of proximal development, Vygotsky, 1978, 84-91), they are in the process of rising into a new level of skills and knowledge and progress depends on the understanding of some detail. When they acquire a hint to solve this problem (scaffold), they experience a break through and can progress to the next step in their studies (Chalmers, 2003).

6 Added Value

A formal learning situation, planned by a teacher or a mentor, can be carried out in many ways, for example through cooperative or individual learning approaches, directed by a teacher or as group work or individual practice. When computers and digital learning material are used in a learning situation, it is expected to introduce definite added value to the learning in comparison to, for example, printed material, and material produced by teacher or the students themselves. The added value is usually in the form of creative use of the possibilities that the computer offers, for example voice, image and video files: the learners can choose a media that best fits their preferences. Jansen, van den Hooven, Jägers and Steenbakkers (2002) point out that especially young students are familiar with computers and multimedia programs (for example so-called video games) and so similar components in learning material suit their life styles and future work. Jansen et al. (2002) have devised a list of aspects of computer-assisted learning that offer added value: (1) adaptability to individual needs, (2) number of flexible options, (3) learning is controlled by the learner, initiated by the learner and is in the form that the learner desires, (4) interesting contents, (5) development of communication, and (6) active participation of the students. In practice, the realization of all the items on this list requires that the developers of the learning material have multidisciplinary skills and knowledge, experience or teaching and time to develop the learning material. The system or learning material should offer the students with tools that are suited to control the contents of the learning material and that make the use of the material more effective and economic. The student should have a feeling that the topic is best learned through the use of a computer.

Table 3. Factor structure of the second pedagogical usability dimension (2. Learner activity)

	Target group					
	Child		Teacher		Adolescent/adult	
Factors	Items LMS	Items LM	Items LMS	Items LM	Items LMS	Items LM
2.1 Reflective thinking		3, 4r				3
2.2 Problem-based learning (PBL)		5, 9		5, 9		5, 9
2.3 Use of primary data sources		6		6		6
2.4 Immersion		7		7		7
2.5 Ownership		8, 10		8, 10		8, 10

Note. LMS = Learning management system. LM = Learning module/material. r = A reversed item, i.e. values in 1 to 5 Likert scale change as follows: 1 -> 5, 2 -> 4, etc. See Appendix for the item labels.

7 Motivation

Motivation (Latin: Movere, to move) affects all learning and makes people behave the way they do. Behaviourists explain the motivation to do things by reference to instincts, desires and reinforcement; cognitive theorists rely on models of cognitive processes and analysis (Wilson & Myers, 2000, 65). Motivations, which can be either consciously or subconsciously goal-oriented, support the direction of an individual's general behaviour (Ruohotie, 1996). Key concepts of motivation include incentives, self-regulation, expectations, attributions of failure and success, performance or learning goals, as well as intrinsic or extrinsic goal orientation (Reeves, 1994; Ruohotie & Nokelainen, 2003). Someone with intrinsic goal orientation strives to reach learning goals for his or her own purposes, because the material is interesting in itself. Someone with an extrinsic goal orientation strives to achieve better results than others (highest grades in the class), to achieve an extrinsic reward (pay raise, grant) or to avoid punishment (for example repeating a course). Contextual motivation, relating, for example, to the interest of the studied topic, varies dynamically. General level motivation is more static, because it can change according to a stage of life. Attitude and motivation can be distinguished, because attitude affects the quality of one's work, while motivation affects one's alertness and vigour (Ruohotie, 1996).

8 Valuation of Previous Knowledge

Learning material that *presumes* previous knowledge from the learner may expect the learner to already possess some skills or knowledge that have been presented, for example, in some earlier learning materials. Learning material that *respects* the learner's previous knowledge takes into account individual differences in skills and knowledge and encourages them to take advantage of it during studies. This approach favours learner's elaboration, contemplation or new issues and the analysis of their relationship with learner's earlier knowledge constructs (Wilson & Myers, 2000). Computer-assisted learning material can include various predefined "paths" that demonstrate the use of the learning material depending on the previous knowledge of the learner. The learning material may review the central concepts from earlier studies that are important for understanding the present material. In this case, the importance of the *learner's* previous knowledge is diminished, but the importance of the previous material and the cumulative nature of knowledge become clear to the learner.

9 Flexibility

Flexible learning material takes into account learners' individual differences. For example, a test given at the beginning of the studies can provide information about previous knowledge, interest towards learning the topic, and expectations of what the learner seeks to gain from the studies. Information gained in pre-testing can be used to provide the learner with optional or alternate routes in the studies (Hannafin & Peck, 1988, 48; Wilson & Myers, 2000). The learner should be given a chance to navigate freely through the learning material. Flexibility in the contents of the learning material means that the material contains diverse assignments (Quinn, 1996). The more adaptable and broadly defined the assignments are (content description with meta coding), the easier it is to combine them to fit the student's individual needs (Leflore, 2000). Collis and Moonen (2002, 10) address the need for social organization of learning (face-to-face, group, individual), language to be used during the learning situations, modality and origin of the learning resources (teacher, student, library, Internet) and instructional organization of learning (assignments, monitoring). They further suggest that students should share the responsibility of identifying appropriate additional learning resources and even contribute to the learning resources (*ibid*, 13).

10 Feedback

The system or learning material should provide the student with encouraging (Albion, 1999; Quinn, 1996) and immediate feedback. Encouraging feedback increases learning motivation; immediate feedback helps the student to understand the problematic parts in his or her learning. Immediate feedback is particularly important in behaviourist (stimulus – reaction) learning materials (Wilson & Myers, 2000). In interaction between a student and a computer, the benefit is that feedback can be given immediately after the student's action. On the other hand, computer provided feedback is rarely so valuable and well timed that it can support learning by itself. If the other side of the dialogue is a human (for example, a peer learner on a discussion forum for the learning material), the feedback is more likely to support reflection, depending, of course, on the quality of interaction inside the peer group (Kolodner & Guzdial, 2000).

In so-called faultless learning, the learning is based on the fact that mistakes are corrected immediately and one cannot progress to the next part in the material before this is done. On the other hand, in a constructivist learning situation, in which the problem is solved in cooperation, the learners may sometimes come up with a faulty solution to an ill-structured problem (Jonassen, Peck & Wilson, 1999, 196). In this case, the teacher or mentor can react to the mistake after a delay and give the students hints at the correct solution or links to additional material.

Empirical Evaluation of the Criteria

A PMLQ (Pedagogically Meaningful Learning Questionnaire) research instrument was developed on the basis of the aforementioned technical and pedagogical usability criteria. The first version of the instrument contained 92 multiple-choice items. The five-point scale ranged from 1 (totally disagree) to 5 (totally agree). The sixth response option was "Not applicable".

The PMLQ has three parts. The first part concerns technical and pedagogical usability of the learning platform (or system) containing 43 items; the second is about technical usability of the learning material (24 items), and the third part measures pedagogical usability of the learning material (25 items). The propositions are clearly marked when measuring issues about system or contents.

Empirical measurements were carried out in October 2003 with 5th and 6th grade elementary school students (n = 66) and their teachers (n = 4). Twenty-four of the children were boys and 42 were girls. Three teachers were male and one was female. Participants evaluated the OPIT learning system and four learning modules embedded into the system. Two of the modules were about mathematical topics (decimal numbers and fractions) and the remaining two modules were about the third foreign language of Finnish school children, English (singular vs. plural and knowledge test about British Isles).

The design of this empirical test contained three stages: First, each participant was profiled using a previously developed ACALQ (Abilities for Computer Assisted Learning Questionnaire) instrument that characterizes respondents by their self-rated motivational level, metacognitive preparedness and social abilities (Nokelainen & Ruohotie, 2004). This information was intended for future purposes in order to control individual differences in the evaluation of learning materials. Second, participants filled out a PMLQ questionnaire for the OPIT platform. Shortly after using each of the four modules they filled out a similar usability questionnaire for each module. Third, for each module, individual score and turnaround time was recorded.

Results of this first version of the PMLQ instrument showed that full scale of 1 to 5 was in use. Dependencies between the variables were investigated with Bayesian dependency modeling (Myllymäki, Silander, Tirri & Uronen, 2002; Nokelainen & Tirri, 2004) due to discrete measurement scale and small sample size. The results supported the chosen dimensionality (Nokelainen, 2004b).

The next step in the development process was to revise the old items because interviews with the children and teachers revealed some deficiencies in wording. The second version of the PMLQ's pedagogical usability section contained 56 items after five new items were included (see Appendix). Empirical measurements were carried out in February 2004 with 5th and 6th grade elementary school students (n = 74) with age median of 12 years. Participants evaluated the OPIT learning system and two learning modules embedded into the system. First module was about decimal numbers, and the second one was about English language singulars and plurals. Both modules were chosen by the teacher and research team. The structure of math module was non-linear, allowing the learner to choose from several equal paths to complete the sequence of numbers tasks. The English module represented typical fill-in test that proceeds in a linear fashion with minimal user control.

Analysis showed that although the distribution of most items was skewed, the full five-point Likert scale was in use.

Table 4 contains the Cronbach's alpha reliability scores for the first part of the study measuring usability of the OPIT learning management system. Fifty-two out of 74 students evaluated the system with the PMLQ. Only a few items measure the pedagogical usability of a learning management system, as our main focus of interest is in the contents (i.e., learning material). As the alpha coefficient tends to grow as the number of items increases, it is not surprising that the third dimension, Cooperative/Collaborative learning, has higher alpha value (.86) than the fourth dimension, Goal orientation (.72).

Table 5 contains the reliability coefficients for the dimensions measuring the pedagogical usability of the two learning modules. Thirty-four students used and evaluated the "Decimal numbers in a continuum" math learning material. Reliability coefficients range from .75 (7. Motivation) to .87 (5. Applicability). Although high alpha value indicates unidimensionality, the inner structures of the dimensions described in this paper are in most cases quite complex, as presented earlier in tables 2 and 3. The second part of the Table 5 shows reliability coefficients for the English "singular or plural" learning material. The alphas are satisfactory ranging from .80 (9. Flexibility) to .92 (5. Applicability).

When we compare the pedagogical usability profiles of the two modules (presented in Table 5), we notice that learner control, activity and flexibility dimensions have higher mean scores for the math module (M = 3.7, SD= .94; M = 3.6, SD= .79; M = 3.4, SD= .66) than for the English module (M = 3.4, SD= .87; M = 3.3, SD= .94; M = 3.2, SD= .68). Although the mean differences are quite small, the measurement instrument is able to capture differences in the pedagogical usability profiles. This leads us to cautiously conclude that the finding supports our earlier hypotheses regarding the different roles of the two learning modules: The math module is designed to allow a learner possibility to complete exercises of a different kind and to decide where and when to proceed, the English module presents more inflexible and linear way of practising singulars and plurals with only one type of exercise.

Description of the Utilization of the Criteria

A typical user of the criteria for pedagogical usability of digital learning material is a teacher, application programmer, or student. The following describes the use of the criteria on a general level:

1. The user logs into the *eValuator* system. S/he submits background and profile information by completing self-assessment questionnaire scales on, for example, motivation, learning strategies and social skills.
2. The user chooses to either search earlier evaluations or to assess new learning material.
3. The user provides criteria for the search or the evaluation (e.g., target group, topic, purpose of the material), defines the target (e.g., LM, ULM or LMS) and, if needed, the usage context of the material (e.g., autonomous studying at school, autonomous studying at home, teacher-directed learning) and the type of material (e.g., supporting/additional assignments, learning a new topic).
4. On the basis of the above criteria, the application generates an evaluation form, which contains statements drawn from the criteria. Each target group has a different set of statements (e.g., children in the 5th or 6th grade who evaluate learning material from the point of view of their own learning, teachers who read evaluations from other teachers or evaluate material that they themselves use, adolescents in polytechnics or

universities who choose material for their own use or evaluate material that they use, and adults in working life).
5. The user provides his or her subjective evaluation of the usability of the material by responding to statements.
6. At the end of search or evaluation, the user is presented with an overview of how the material corresponds with the different parts of the technical and pedagogical usability criteria, as evaluated by the user.
7. The system serves also as a search tool. When the evaluation database contains more evaluations, user is able to conduct personalised searches in order to find the most appropriate learning material. The personal profile (collected with ACALQ and various background questions) and pedagogical usability scores (collected with PMLQ) act as filters. For example, user is able to ask for 5th grade elementary school math LM's that have something to do with decimal numbers and are suitable for collaborative working. A second example might be to ask for such university level statistical LM's that are rated high by registered users whose personal profile is close to the users' profile.

Table 4. Reliability coefficients of the dimensions measuring usability of a learning management system

		Sample 1[a]		
Pedagogical Usability Dimensions	Items	*M*	*SD*	α
1. Learner control	—			
2. Learner activity	—			
3. Cooperative/Collaborative learning	11, 12, 13, 14, 15	3.4	1.22	.86
4. Goal orientation	23, 24	3.7	1.11	.72
5. Applicability	—			
6. Added value	35	3.6	1.36	—
7. Motivation	—			
8. Valuation of previous knowledge	—			
9. Flexibility	—			
10. Feedback	52	3.7	1.30	—

[a] Elementary school students, n = 52, age range 10 – 13 years, age median 12 years.
Note. See Appendix for the item labels.

Conclusion

A rapidly growing quantity of digital learning material is available to fill various learning needs via online (e.g., Internet, mobile networks) and offline (e.g., DVD, CD-ROM) media. A broad array of learning material is offered to learners all over the world. However, whether the application selected achieves the desired ends depend on various factors; for example, whether the learner needs to learn new things or just refresh old knowledge, is working alone or collaboratively with other learners, or is willing to use commercial or non-commercial material.

This paper presents criteria for evaluating the pedagogical usability of digital learning material. In practice the role of the criteria is to give the learner a chance to choose the most suitable learning material possible for any learning situation. Previously constructed criteria for evaluating pedagogical usability have been introduced, and differences between them have been considered. The various phases of creating criteria in this study have been described and the empirical research setting has been evaluated. The 56 –item multiple-choice inventory (PMLQ) that operationalises the pedagogical usability criteria was presented. Results of a two-stage empirical evaluation of the criteria with 5th and 6th grade elementary school children (2003 n = 66, 2004 n = 74) was reported. Results supported all theoretical dimensions of the pedagogical usability criteria. The PMLQ was able to capture differences in the pedagogical usability profiles of the learning modules.

Table 5. Reliability coefficients of the dimensions measuring usability of two learning modules (math and English)

		Math Sample 2[a]			English Sample 3[b]		
Pedagogical Usability Dimensions	Items	*M*	*SD*	α	*M*	*SD*	α
1. Learner control	1, 2, 47, 48, 49, 40	3.7	.94	.86	3.4	.87	.90
2. Learner activity	3, 4, 5, 6, 7, 8, 9, 10	3.6	.79	.86	3.3	.94	.91
3. Cooperative/Collaborative learning	11, 12, 13, 14, 15	3.4	.90	.76	3.4	1.05	.88

4. Goal orientation	20, 21, 22, 23, 24	3.5	.91	.78	3.3	.97	.85
5. Applicability	25, 26, 27, 28r, 29, 30, 31, 54, 55, 56, 40, 48, 49	3.5	.75	.87	3.4	.85	.92
6. Added value	32, 33, 34, 35, 45	3.1	1.05	.82	3.2	1.10	.88
7. Motivation	13, 36, 37, 38	3.6	.90	.75	3.4	1.03	.83
8. Valuation of previous knowledge	39, 40, 41	3.4	.96	.79	3.3	.97	.85
9. Flexibility	42, 44, 46	3.4	.66	.77	3.2	.68	.80
10. Feedback	50, 51, 52, 53	3.7	.76	.79	3.6	.97	.89

[a] Elementary school students, n = 34, age range 10 – 13 years, age median 12 years.
[b] Elementary school students, n = 74, age range 10 – 13 years, age median 12 years.
Note. r = A reversed item, i.e. values in 1 to 5 Likert scale change as follows: 1 -> 5, 2 -> 4, etc. See Appendix for the item labels.

Discussion

The first drawback of the pedagogical usability criteria presented here is their generalisability to other domains. In this study, generalisability is limited by various factors, for example, the small sample size, the narrow age range of respondents' and the limited number of learning material evaluated. To address these issues, we are currently conducting empirical studies that aim at evaluating more representative set of learning material in different domains targeted for both adolescent and adult learners. The criteria of pedagogical usability will in the future be complemented by a segment evaluating "mobile usability" (Syvänen, Nokelainen, Ahonen & Turunen, 2003).

The second drawback, although in a more debatable sense, is the methodology we have applied to measure usability, that is, a self-report questionnaire aiming to measure subjective end-user satisfaction. According to researchers specialising in the measurement of usability (Rubin, 1994; Nielsen, 1993; Kirakowski, 2003), the use of questionnaires is normally justified in usability research as part of a test arrangement where a user answers a pre-test questionnaire, uses the application that is to be evaluated and then answers a post-test questionnaire evaluating the usability. However, Kirakowski (2003) sees self-evaluation questionnaires as particularly appropriate for measuring subjective matters (e.g., the enjoyability of use) in usability research. We all agree, that various other test arrangements (Rubin, 1994; Nielsen, 1993, 192-200) are better for gathering factual data (e.g., time spent in doing a particular task). On the other hand, the criteria developed here can also be used as the basis for heuristic evaluation (Nielsen & Molich, 1990).

The third drawback is that the criteria do not directly address the issue of cultural sensitivity (Reeves, 1994). Authority (e.g., who is 'permitted' to ask questions or 'obligated' to lead group work on the basis of his/her social status) and collaboration (e.g., are questions regarding the peer-learners presentation or evaluation of other learners' work allowed) are examples of cultural issues that are strongly connected to pedagogics. The problem of describing cultural sensitivity via pedagogical criteria is two-fold. First, it is to some extent connected with the metadata (the definition of creator, ideal target group, religion, language and country) and technical usability (graphical user interface colours and symbols) issues. Second, cultural sensitivity is very strongly latent variable by nature, and hence it is not trivially operationalised into questions that end-user is able to answer.

Finally, if we try to verbalise what the pedagogical usability criteria aim to address, we may ask: "Does the system, and/or learning material it contains, make it possible for the student and the teacher to achieve their goals?" The important thing to acknowledge is that the dimensions of technical and pedagogical usability (operationalised dimensions) might correlate and thus describe (i.e., in factor analysis language "load on") both usability and utility (latent dimensions). From the perspective of the user/learner, the number of things to be memorised and the feedback required are examples of these kinds of common components between the user and the system, and between the user and the contents of the learning material. From the viewpoint of usability, the system works well in user terms when the user does not have to resort to any help and is able to control the program, and not the other way around. The number of things to be memorised is, in terms of utility, related to the functioning of the learning material because this will be minimized if the studied subject matter is presented in manageable sections that make sense to the users. There may be a certain degree of tension between some of the criteria. For example, usability is enhanced if the number of things to be remembered and errors are kept to a minimum, but for the utility of the learning material it is desirable for the user to remember the core set of functions so that they can be used readily without prompting. Another example is authenticity and usability. Using real representations from the outside world in the user interface may not be as efficient as a metaphoric design. In the design of learning materials, the aim is often to correspond closely with the real world so that it will be easier for users to link new knowledge to earlier knowledge structures. On the other hand, learning materials intended to inculcate a certain skill by repetition can be weak in their real-world correspondence but still produce good learning results, particularly if the skill being learned is not derived directly from the real world, either.

The criteria described here are not the first attempt to provide an analysis tool for the evaluation of the usability of digital learning material; both its structure and content are to great extend based on previous research. However, the criteria and research that led to its creation, have the following original features that bring added value for research on the pedagogical usability of digital learning material: 1) a multidisciplinary approach involving the fields of both computer technology and education, 2) a review of previous research, 3) the development of the criteria on the basis of systematic empirical research, performed with the help of real users; and 4) a multiple choice inventory that was developed on the basis of the criteria and can be used independently or as part of the eValuator computer application. A preliminary version of the eValuator online software is available for educational and research purposes at http://evaluator.hamk.fi.

Acknowledgements

This study is part of Digital Learning research project (http://dll.hamk.fi/dl2/en/) financed by National Technology Agency in Finland (http://www.tekes.fi/eng/), Research Centre for Vocational Education, University of Tampere, Finland (http://rcve.uta.fi/), and Hame Polytechnic, Finland (http://www.hamk.fi/in-english/). Author wishes to thank Mikko Horila, Miikka Miettinen, Dr. Jorma Saarinen, Antti Syvänen, Tuomo Tammi and Jan Överlund for their contribution to this project. Dr. Bruce Beairsto's comments helped to improve the quality of this manuscript.

References

Albion, P. (1999). Heuristic evaluation of educational multimedia: From theory to practice. *Paper presented at the 16th Annual conference of the Australasian Society for Computers in Learning in Tertiary Education*, retrieved November 24, 2005 from http://www.usq.edu.au/users/albion/papers/ascilite99.html.

Barab, S. A., & Duffy, T. M. (2000). From Practice Fields to Communities of Practice. In Jonassen, D. H. & Land, S. M. (Eds.), *Theoretical Foundations of Learning Environments*, Mahwah, NJ, USA: Lawrence Erlbaum Associates, 25-56.

Chalmers, P. (2003). The role of cognitive theory in human -computer interface. *Computers in Human Behavior, 19*, 593-607.

Chin, J., Diehl, V., & Norman, K. (1988). Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface. *Proceedings of ACM CHI'88 Conference on Human Factors in Computing Systems*, New York: ACM Press, 213-218.

Collis, B., & Moonen, J. (2002). *Flexible Learning in a Digital World. Experiences and expectations*, London: Kogan Page.

Forman, E. A., & Cazden, C. B. (1985). Exploring Vygotskian perspectives in education: The cognitive value of peer interaction. In Wertsch, J. (Ed.), *Culture, communication, and cognition: Vygotskian perspectives*, Cambridge: Cambridge University Press, 323-347.

Hannafin, M., & Peck, K. (1988). *The Design, Development and Evaluation of Instructional Software*, New York: Macmillan.

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

Jansen, W., van den Hooven, H., Jägers, H., & Steenbakkers, G. (2002). The Added Value of E-learning. *Paper presented at the Informing Science & IT Education*, June 19-21, 2002, Cork, Ireland.

Jonassen, D. H. (1995). Supporting Communities of Learners with Technology: a Vision for Integrating Technology with Learning in Schools. *Educational Technology, 35* (4), 60-63.

Jonassen, D. H., Myers, J., & McKillop, A. (1996). From Constructivism to Constructionism: Learning with Hypermedia/Multimedia Rather Than from It. In Wilson, B. (Ed.), *Constructivist Learning Environments,* Englewood Cliffs, NJ, USA: Educational Technology Publishers, 93-106.

Jonassen, D. H., Peck, K., & Wilson, B. (1999). *Learning with Technology. A Constructivist Perspective*, Upper Saddle River, NJ, USA: Merrill.

Keinonen, T. (1998). *One-dimensional usability – Influence of usability on consumers' product preference*, Saarijärvi, Finland: Gummerus.

Kirakowski, J. (2003). *Questionnaires in Usability Engineering*, retrieved November 24, 2005 from: http://www.ucc.ie/hfrg/resources/qfaq1.html.

Kolodner, J., & Guzdial, M. (2000). Theory and Practice of Case-Based Learning Aids. In Jonassen, D. H. & Land, S. (Eds.), *Theoretical Foundations of Learning Environments,* Mahwah, NJ, USA: Lawrence Erlbaum, 215–242.

Kurhila, J., Miettinen, M., Nokelainen, P., & Tirri, H. (2002). Use of Social Navigation Features in Collaborative E-Learning. Paper presented at the E-Learn 2002 Conference, retrieved May 1, 2006 from http://cosco.hiit.fi/edutech/publications/elearn2002.pdf.

Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*, New York, USA: Cambridge University Press.

Leflore, D. (2000). Theory Supporting Design Guidelines for Web-Based Instruction. In Abbey, B. (Ed.), *Instructional and Cognitive Impacts of Web-Based Education*, Hershey: Idea Group Publishing, 102-117.

Lin, H., Choong, Y., & Salvendy, G. (1997). Proposed Index of Usability: A Method for Comparing the Relative Usability of Different Software Systems. *Behaviour and Information Technology, 16* (4/5), 267-278.

Littleton, K., & Häkkinen, P. (1999). Learning Together: Understanding the Processes of Computer-Based Collaborative Learning. In Dillenbourg, P. (Ed.), *Collaborative Learning. Cognitive and Computational Approaches*, Amsterdam: Pergamon, 20-30.

Livingstone, D. W. (2000). *Exploring the Icebergs of Adult Learning: Findings of the First Canadian Survey of Informal Learning Practices*, NALL Working Paper #10-2000, Ontario, Canada: Ontario Institute for Studies in Education, University of Toronto.

Mayes, J., & Fowler, C. (1999). Learning technology and usability: a framework for understanding courseware. *Interacting with Computers, 11*, 485-497.

Miller, G. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review, 63*, 81-97, retrieved November 24, 2005 from: http://www.well.com/user/smalin/miller.html.

Myllymäki, P., Silander, T., Tirri, H., & Uronen, P. (2002). B-Course: A Web-Based Tool for Bayesian and Causal Data Analysis. *International Journal on Artificial Intelligence Tools*, 11 (3), 369-387.

Nielsen, J. (1990). Evaluating Hypertext Usability. In Jonassen, D. H. & Mandl, H. (Eds.), *Designing Hypermedia for Learning*, Berlin: Springer-Verlag, 147-168.

Nielsen, J. (1993). *Usability Engineering*, Boston: Academic press.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. *Proceedings of the ACM CHI'90 Conference,* New York, USA: ACM Press, 249-256.

Nokelainen, P. (2004a). Designing an Evaluation Tool for Digital Learning Materials. *Paper presented at the Society for Information Technology and Teacher Education International Conference*, March 1-6, 2004, Atlanta, USA.

Nokelainen, P. (2004b). Conceptual Definition of the Technical and Pedagogical Usability Criteria for Digital Learning Material. *Paper presented at the EdMedia 2004 Conference*, June 21-26, 2004, Lugano, Switzerland.

Nokelainen, P., & Ruohotie, P. (2004). Empirical Validation of Abilities for Computer Assisted Learning Questionnaire. In Marsh, H. W., Baumert, J., Richards, G. E. & Trautwein, U. (Eds.), *Proceedings of the 3rd International Self-Concept Research Conference,* Sydney: University of Western Sydney, Self Research Centre, 676-689.

Nokelainen, P., & Tirri, H. (2004). Bayesian Methods that Optimize Cross-cultural Data Analysis. In Campbell, J. R., Tirri, K., Ruohotie, P. & Walberg, H. (Eds.) *Cross-cultural Research: Basic Issues, Dilemmas, and Strategies*, Hämeenlinna, Finland: RCVE, 141-158.

Palloff, R. M., & Pratt, K. (2005). *Collaborating Online. Learning Together in Community*, San Francisco, CA, USA: Jossey-Bass.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction Design: Beyond Human-Computer Interaction*, New York, NY, USA: John Wiley & Sons.

Quinn, C. (1996). *Pragmatic Evaluation: Lessons from Usability*, retrieved November 24, 2005 from: http://www.ascilite.org.au/conferences/adelaide96/papers/18.html

Reeves, T. C. (1994). Evaluating what really matters in computer-based education. In Wild, M. & Kirkpatrick, D. (Eds.) *Computer education: New Perspectives*, Perth, Australia: MASTEC, 219-246.

Rubin, J. (1994). *Handbook of Usability Testing: How to plan, design, and conduct effective tests*, New York, USA: John Wiley & Sons.

Ruohotie, P. (1996). Professional Growth and Development. In Leithwood, K. (Ed.), *International Handbook of Educational Leadership and Administration*, Dordrecht: Kluwer Academic Publishers, 419-445.

Ruohotie, P., & Nokelainen, P. (2003). Practical Considerations of Motivation and Computer-supported Collaborative Learning. In Varis, T., Utsumi, T. & Klemm, W. R. (Eds.), *Global Peace Through The Global University System*, Hämeenlinna, Finland: RCVE, 226-236.

Savery, J. R., & Duffy, T. M. (1995). Problem based learning: An instructional model and its constructivist framework. *Educational Technology, 35* (5), 31-7.

Shneiderman, B. (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (3rd Ed.), Menlo Park, CA, USA. Addison Wesley.

Squires D., & Preece, J. (1996). Usability and Learning: Evaluating the Potential of Educational Software. *Computers & Education, 27* (1), 15-22.

Squires, D., & Preece, J. (1999). Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. *Interacting with Computers, 11*, 467-483.

Syvänen, A., Nokelainen, P., Ahonen, M., & Turunen, H. (2003). Approaches to Assessing Mobile Learning Components. *Paper presented at 10th Biennial Conference of the European Association for Research on Learning and Instruction*, August 26-30, 2003, Padova, Italy.

Tognazzini, B. (2003). *First principles*, retrieved November 24, 2005 from: http://www.asktog.com/basics/firstPrinciples.html.

Tudge, J., & Rogoff, B. (1989). Peer influences on cognitive development: Piagetian and Vygotskian perspectives. In Bornstein, M. H. & Bruner, J. H. (Eds.), *Interaction in human development*, Hillsdale, NJ, USA: Lawrence Erlbaum, 17-40.

Vygotsky, L. S. (1978). *Mind in Society. The Development of Higher Psychological Processes*, London, England: Harvard University Press.

Wilson, B., & Myers, K. (2000). Situated Cognition in Theoretical and Practical Context. In Jonassen, D. H. & Land, S. (Eds.), *Theoretical Foundations of Learning Environments*, Mahwah, NJ, USA: Lawrence Erlbaum Associates, 57-88.

Appendix 1

The pedagogical usability items for 4 to 6th grade elementary school students

1. When I worked on this assignment, I felt that I, not the program, held the responsibility for my own learning. (Definition: I'm not repeating same kind of steps all the time in my studying, but the assignments make me think and make a different solution for each one.)
2. When I used this learning material, I felt that I controlled what it did and not vice versa. (Definition: The program does not lead me from one step to another, but I can choose by myself in which order I want to finish my tasks.)
3. I have to think and make my own solutions to learn this learning material. (Definition: I have to concentrate on this material, I cannot complete the tasks simply with role learning.)
4. This learning material has been divided into sections, my task is to learn them in pre-defined order (and possibly respond to assignments).
5. This learning material provides learning problems without a pre-defined model for the solution.
6. This material does not have material in itself, but links to various other sources, which I have to use to learn. (Definition: If the topic is ”a dumb yard”, there is no ready-made presentation of the topic. The learning starts, for example, with a short tip by the teacher, which tells what kinds of things are needed in order to build a dumb yard. Some of the information that you need may be in the system, but you have to find most of it from newspapers, books or the Internet in order to make your own presentation.)
7. I get so deep into this learning material that I forget all about what is happening around me and how much time I spend on it.
8. When I work with this learning material, I feel that I know more about some topics than others, I'm ”an expert”. (Definition: The learning material may involve an individual information gathering task, for example, an interview of neighbours or measuring the depth of packed snow in one's home garden over the period of one month.)
9. When I work in this learning material, I (or us, if a group work) have to find out own solutions without the teacher's or the program's model solutions.
10. I am proud of my own solution, or one that I made with others, to the problem presented in the learning material. (Definition: I feel that I, or we together, have made something that is significant.)
11. This learning material lets me talk with my classmates. (Definition: For example, messages in chat or notice board.)
12. I can do group work with my classmates in this learning material. (Definition: If I wanted, I could do assignments together with my classmate so that we both used our own computers.)
13. It is pleasant to use the learning material with another student on the same computer.
14. This learning material lets me know what other users have been doing in the system. (Definition: For example, which learning materials have been read the most or assignment that have been the most popular.)
15. This learning material lets me know what other users are doing when I'm using the system. (Definition: For example, the most read material at the moment or the assignment with which most people are working on.)
16. This learning material offers me simple utility programs (for example, a calculator).
17. This learning material offers me versatile utility programs (for example, Excel sheets, a HTML editor, text processor, etc.).
18. In this learning material the utility programs have a central role. (Definition: I have to, for example, edit an Excel sheet to solve a problem.)
19. I can save my work on this learning material and use or evaluate others' work. (Definition: I can, for example, explore or evaluate other groups' group works and use them in my own studies.)
20. This learning material tells me clearly what I'm expected to know (or learn) after I've used it. (Definition: The learning goals are clearly set, for example, ”After this assignment, you will know how to divide with decimal fraction” or ”After this assignment you can form interrogative clauses in English”.)
21. This learning material tells me clearly why it is useful for me to learn this material. (Definition: The learning goals are justified, for example, ”This assignment will help you to make interrogative phrases in English.”)
22. The learning material assesses my achievements with scores. (Definition: For example, the system gives a score at the and of an assignment and shows the maximum score.)
23. This learning material tells me how much progress I have made in my studies. (Definition: I know what I have practiced or learned thus far.)
24. This learning material is strictly limited. (Definition: For example, the topic of a math learning material is “Calculating the mean”.)
25. This learning material teaches me skills that I will need. (Definition: I will be able to, for example, convert euros into crowns or marks, or help my parents to choose between different-sized packages according to their prize difference.)

26. I feel that I will be able to use the skills and knowledge this learning material has taught me in the future.
27. This learning material is based on the idea that ”one learns the best by doing stuff by oneself”. (Definition: The material offers more assignments than, for example, PowerPoint presentations.)
28. I feel that this learning material will help me to do better in the test. (Definition: I think the assignments in the material are similar to the assignment that we usually have in tests.)
29. This learning material is suitably challenging to me. (Definition: The assignments are not too easy or too hard.)
30. I feel that this learning material has been designed for me. (Definition: The material suits your own needs, and it does not feel that you are considered too smart or too dumb.)
31. This learning material adjusts the difficulty to suit my skills. (Definition: I can practice something that is hard for me until I have learned it and before I move on to the next topic.)
32. The images in this learning material help me to learn.
33. The sounds in this learning material help me to learn.
34. The animations in this learning material help me to learn.
35. It is more useful to me to learn topics with this learning material than with conventional methods in a classroom. (Definition: Think if you would be more willing to do this assignment with a computer or with a normal study book or exercise book.)
36. I try to achieve as high a score as I can in this learning material.
37. I want to learn the topics of this learning material as deeply as I can.
38. I’m interested in the topic of this learning material.
39. This learning material required me to know something that had been taught in some other learning material. (Definition: This material made a reference to some other learning material.)
40. I can use my earlier knowledge when I study with this material.
41. This learning material goes over earlier material before starting to teach a new topic. (Definition: For example, in mathematics, the material first goes over simpler calculations that are needed to learn a more difficult topic.)
42. This learning material offers optional routes for my progress. (Definition: I can choose different assignments each time I use the system.)
43. This learning material does not let me proceed to the next point before I have answered correctly to every question. (Definition: For example, in an English language assignment one has to answer correctly to every question, even with the help of the program, before it lets you proceed to the next topic.)
44. This learning material has many similar, consecutive assignments. (Definition: For example, an English fill-in assignment that has many consecutive assignments for am, are, and is sentences.)
45. This learning material makes it quick and easy for me to learn a new topic or recap an earlier topic.
46. If I cannot remember a particular word or concept while using this learning material, I can go back and check its meaning in previous material.
47. When I used this learning material, I felt that I had to remember too many things at the same time. (Definition: I felt at some points that I should have used a paper to write some things down.)
48. This learning material presents information in a format that makes it easy to learn. (Definition: Information is presented in meaningful, interconnected entities, and not in separate pieces that are hard to understand.)
49. This learning material presents new material (or recaps old) in ”portions” suitable for me. (Definition: There are not too many new things presented at once, I have time to learn them before moving onto the next topic.)
50. I can make a certain number of mistakes with this material (for example, wrong answers to calculus tasks), after which the program shows me the correct answer.
51. When I make a wrong solution in an assignment, the program gives me a friendly note.
52. This learning material gives me motivating feedback. (Definition: I am willing to try out the less used functions in the learning material, because I know that the system will give me all the advice that I need.)
53. This learning material provides me with immediate feedback of my activities. (Definition: When I write my response to a calculus task, the system shows me immediately whether the answer is correct or not.)
54. This learning material gives first an example of the correct solution. (Definition: Multiplying with decimal fractions is started with a model performance, after which I will calculate on my own.)
55. In this learning material, I get to carry the responsibility for the solution of an assignment in small portions. (Definition: For example, in a math task, I will be first shown the task and then the result. Next, I see the task but not the result, which I have to solve on my own.)
56. I think I learn more quickly with this material than normally. (Definition: This learning material provides me with the right kind of support when I need it.)

Appendix 2

The pedagogical usability subdimensions

Note. Related items are presented in parenthesis (see Appendix 1). r = rewersed, i.e., negatively worded, item.

1. Learner control
1.1 Minimum memory load (47)
1.2 Meaningful encoding (48,49)
1.3 Take responsibility for one's own learning (1)
1.4 User control (2)
8.2 Elaboration (40)

2. Learner activity
2.1 Reflective thinking (3, 4r)
2.2 Problem-based learning (5, 9)
2.3 Use of primary data sources (6)
2.4 Immersion (7)
2.5 Ownership (8, 10)
2.6 Primary data source (for PBL) (Items for teacher only)
2.7 Facilitative teacher (Items for teacher only)
2.8 Didactic teacher (Items for teacher only)
2.9 Individual/distance learning (Items for teacher only)

3. Cooperative/Collaborative learning
3.1 Support for conversation and dialog (11)
3.2 Group work (12,13)
3.3 Asynchronous social navigation (14)
3.4 Synchronous social navigation (15)
3.5 Asynchronous social navigation monitoring (Items for teacher only)
3.6 Synchronous social navigation monitoring (Items for teacher only)
3.7 Tertiary courseware (19)

4. Goal orientation
4.1 Explicit goals (20)
4.2 Usefulness of goals (21)
4.3 Focus on results (22)
4.4 Focused goals (34)
4.5 Monitor one's own studies (pedagogic feedback) (23)
4.6 Set one's own goals (Items for teacher only)

5. Applicability
5.1 Authentic material (25,28r)
5.2 Perceived usefulness (26)
5.3 Learning by doing (27)
5.4 Adequate material for the learners needs (human development) (29)
5.5 Pretesting and diagnostics (30,31)
5.6 Prompting (54)
5.7 Fading (55)
5.8 Scaffolding (56)
1.2 Meaningful encoding (40,48,49)

6. Added value
6.1 Overall added value for learning (35)
6.2 Effectiveness for learning (45)
6.3 Added value of pictures (32)
6.4 Added value of sounds (33)
6.5 Added value of animations (34)

7. Motivation
7.1 Intrinsic goal orientation (37)
7.2 Extrinsic goal orientation (36)
7.3 Meaningfulness of studies (38)
2.4 Immersion (7)

8. Valuation of previous knowledge
8.1 Prerequisites (39)
8.2 Elaboration (40)
8.3 Examples (41)

9. Flexibility
9.1 Pretesting and diagnostics (42)
9.2 Task decomposition (43r,46)
9.3 Repetitive tasks (44)

10. Feedback
10.1 Encouraging feedback (52,51)
10.2 Accurate feedback (53)
10.3 Errorless learning (50)