Multiple Criteria Comparative Evaluation of E-Learning Systems and Components

Eugenijus KURILOVAS, Valentina DAGIENĖ
Institute of Mathematics and Informatics
Akademijos 4, LT-08663 Vilnius, Lithuania
e-mail: eugenijus.kurilovas@itc.smm.lt; dagiene@ktl.mii.lt

Abstract. The main scientific problems investigated in this paper deal with the problem of multiple criteria evaluation of the quality of the main components of e-learning systems, i.e., learning objects (LOs) and virtual learning environments (VLEs). The aim of the paper is to analyse the existing LO and VLE quality evaluation methods, and to create more comprehensive methods based on learning individualisation approach. LOs and VLEs quality evaluation criteria are further investigated as the optimisation parameters and several optimisation methods are explored to be applied. Application of the experts' additive utility function using evaluation criteria ratings and their weights is explored in more detail. These new elements make the given work distinct from all the other earlier works in the area.

Keywords: multiple criteria evaluation of quality, learning objects, virtual learning environments, score-rating, weights, optimisation.

Introduction

Researchers usually divide e-learning systems components (such as learning objects (LOs) and virtual learning environments (VLEs)) quality evaluation criteria into technological, pedagogical and organisational evaluation criteria (Kurilovas, 2005; Kurilovas and Dagiene, 2008).

LO is referred here as "any digital resource that can be reused to support learning" (Wiley, 2000). VLE is considered here as the specific information system which provides the possibility to create and use different learning scenarios and methods (Institute of Mathematics and Informatics, 2005).

In ISO/IEC 14598-1:1999 quality evaluation is defined as "the systematic examination of the extent to which an entity (part, product, service or organisation) is capable of meeting specified requirements" (ISO/IEC 14598-1:1999, 1999).

We can evaluate both internal quality and quality in use of the components of e-learning systems such as LOs and VLEs. According to Gasperovic and Caplinskas (2006), internal quality is a descriptive characteristic that describes the quality of software independently from any particular context of its use, while quality in use is evaluative characteristic of software obtained by making a judgment based on criteria that determine the

worthiness of software for a particular project. Gasperovic and Caplinskas (2006) consider that it is impossible to evaluate quality in use without knowing characteristics of internal quality.

Different scientific methods are used for quality evaluation of software.

The paper is aimed to consider the problems of the expert evaluation of the technological quality of LOs and VLEs.

Expert evaluation is referred here as the multiple criteria evaluation of the e-learning system components aimed at selection of the best alternative based on score-ranking results.

According to Dzemyda and Šaltenis (1994), if the set of decision alternatives is assumed to be predefined, fixed and finite, then the decision problem is to choose the optimal alternative or, maybe, to rank them. But usually the experts (decision makers) have to deal with the problem of optimal decision in the multiple criteria situation where the objectives are often conflicting. In this case, according to Dzemyda and Šaltenis (1994), an optimal decision is the one that maximises the decision maker's utility.

E-learning systems' and components' multiple criteria evaluation method is referred here as the experts additive utility function including the alternatives' evaluation criteria, their ratings (values) and weights. The different components of the e-learning systems (LOs and VLEs in our case) have the different quality evaluation criteria, but it is possible to use the same ratings (values) and weights in their multiple criteria evaluation methods.

The paper is organised as follows: Section 1 analyses LOs evaluation criteria, Section 2 – VLEs evaluation criteria, Section 3 – VLEs multiple criteria evaluation methods using the evaluation criteria ratings (values) and weights, Section 4 – overall conclusions and results.

1. Quality Evaluation Criteria of Learning Objects

1.1. *Different Approaches to Quality Evaluation of Learning Objects*

The provision of LOs provides better access to quality LOs and supports enhanced learning outcomes.

According to Haughey and Muirhead (2005), the purpose of LOs is to increase the effectiveness of learning by making content more readily available, by reducing the cost and effort to produce quality content, and by allowing content to be more easily shared. These two purposes, effectiveness and efficiency, receive differing emphases from different sectors.

One of the main criteria for achieving the high LOs effectiveness and efficiency level is LOs reusability (Kurilovas, 2009). The need for reusability of LOs has at least three elements:

1. Interoperability: LO is interoperable and can be used in different platforms.
2. Flexibility in terms of pedagogic situations: LO can fit into a variety of pedagogic situations.

3. Modifiability to suit a particular teacher's or student's needs: LO can be made more appropriate to a pedagogic situation by modifying it to suit a particular teacher's or student's needs (McCormick *et al.*, 2004).

Reusability criteria are ones of the most important LOs evaluation criteria for the users (Kurilovas, 2009). These characteristics deal with LOs 'internal' quality, but mostly reflect LOs 'quality in use'.

The e-learning system's flexibility and learning individualisation possibilities for the users depend on the level of separating of system 'Content' LOs from 'Activity' ones (so-called 'pedagogical decontextualisation') and their reusability level (Dagiene and Kurilovas, 2008, 2007; Kurilovas and Kubilinskiene, 2008a, 2008b). The flexibility of e-learning system and learning personalisation possibilities for its users are achieved also by separating LO metadata repository from LO repositories which can be on different servers, usage of highly adaptable (open source) software to create and/or reuse LOs, and provide the modular highly adaptable online learning environments – VLEs (Kurilovas and Dagiene, 2008).

This kind of flexible e-learning system design seems to be one of the best possible e-learning solutions from technological, educational, organisational and socio-economic points of view. The detailed evidence of this statement is out of the scope of the article, but the aforementioned approach ensures the system's pedagogical and organisational flexibility (e.g., high learning personalisation level) as well as the better financial and economic efficiency indicators such as less investment into LOs for a single user, major financial benefit, less time to buy off, etc. (Kurilovas, 2007).

This flexibility could be developed because (Kurilovas, 2007):
- Major reusability of main system components is achieved.
- More users can benefit from such system.
- 'Content' and 'Activity' (learning design) creators have the possibility not to reinvent the wheel but to use and improve already created LOs.
- Better conditions are created for various content and/or design creators to improve the quality of existing LOs by their permanent (collaborative) modification.

The evaluation of LOs is a comparatively new concern as the quantity of LOs has grown and the development of LO repositories has come about to allow for greater ease in finding and using LOs for blended learning (Kurilovas, 2009).

According to Haughey and Muirhead (2005), "the growth in the number of LOs, the multiplicity of authors, their increasing diversity of design and their availability to trained and untrained educators has generated interest in how to evaluate them and which criteria to use to make judgments about their quality and usefulness".

1.2. *LORI Quality Criteria*

The need to evaluate LOs requires the development of criteria to be used in judging them. Vargo *et al.* (2003) developed a Learning Object Review Instrument (LORI) to evaluate LOs.

The LORI approach uses the following 10 criteria when examining LOs:
1. Presentation: Aesthetics.

  2. Presentation: Design for learning.

  3. Accuracy of content.

  4. Support for learning goals.

  5. Motivation.

  6. Interaction: Usability.

  7. Interaction: Feedback and adaptation.

  8. Reusability.

  9. Metadata and interoperability compliance.

10. Accessibility.

The criteria 'Reusability', 'Metadata and interoperability compliance' and 'Accessibility' are suitable for the technological evaluation (Kurilovas, 2009).

The LORI criteria were drawn from a review of pertinent literature on instructional design, computer science, multimedia development and educational psychology.

Each measure was weighted equally and was rated on a four point scale from 'weak' to 'moderate' to 'strong' to 'perfect'.

The LORI process involved both individual and group rating of LOs (Vargo *et al.*, 2003).

1.3. *Paulsson and Naeve Quality Criteria*

Six areas for establishing LO technological quality criteria are suggested by Paulsson and Naeve (2006):

  1. A narrow definition.

  2. A mapping taxonomy.

  3. More extensive standards.

  4. Best practice for use of existing standards.

  5. Architecture models.

  6. The separation of pedagogy from the supporting technology of LOs.

According to Paulsson and Naeve (2006), most LO implementations do not nearly meet this vision. For those reasons it is essential to establish common criteria of quality for LOs. Technical quality criteria are specific characteristics and properties that LOs must (or in some cases ought to) adhere to – including best practice, guidelines and standard specifications – in order to be regarded as LOs.

Paulsson and Naeve (2006) have focused on technological quality criteria for LOs. Other quality criteria, such as pedagogical quality, usability or functional quality were beyond the scope of their study. Such aspects of quality are addressed by Van Assche and Vuorikari (2006), where the authors suggest a quality framework for the whole life cycle of LOs.

Evaluation conducted by Paulsson and Naeve focuses on:

  1. Architecture – in terms of separation of data, logics, presentation, and implementation of interaction interfaces.

  2. Pedagogical contextualisation.

  3. The use of standards and the extent to which they are decomposable or composable.

According to Paulsson and Naeve (2006), "many of those issues are directly or indirectly related to the lack of explicit definitions and clear architectural models, together with technical (as well as other) quality criteria that are directly related to technical architecture. Many of the pedagogical dependencies and shortcomings seem to be caused by technical bindings of content to presentation and application logics as well as built in instructional design elements".

The study of Paulsson and Naeve (2006) has shown that there is a huge discrepancy between different definitions of the LO concept. This makes it hard (if not impossible) to develop LOs that have the qualities that LOs are often ascribed in terms of reusability, interoperability, and context independence. Definitions really range from 'anything to everything' (McGreal, 2004). 'Anything to everything' from a 'content' perspective is a good thing as this makes it easier to support different pedagogical directions and methods, but 'anything to everything' from a technological perspective becomes unmanageable. Paulsson and Naeve suggest the technical and pedagogical definitions of LOs to be separated – within a common definition of LOs.

The lack of common low-level definitions and models is a threat to interoperability, technological quality as well as for the acceptance of the LO concept itself.

Paulsson and Naeve (2006) have investigated that the pedagogical content is often of good quality and that the ambitions are set high, but that LOs still do not live up to the expectations that would make them context independent, reusable objects. One important reason is that little consideration is given to fundamental software design principles, such as layering, object orientation, structuring of data etc., which could enhance features usually ascribed to LOs.

To address the identified problems Paulsson and Naeve (2006) suggest six areas for action in order to establish technological quality criteria for LOs:

1. There is a need for a common (more narrow) definition of what is, and what is not a LO.
2. In connection to narrowing down the definition, there is a need for a taxonomy that is reflected in the definition where granularities as well as special properties are regarded.
3. Standards used for LOs should be extended to go beyond descriptive information, such as metadata, sequencing, and packaging to also embrace standards for interfaces, 'machine readable' descriptions of technical properties and interaction interfaces.
4. There is a need to establish standards and recommendations that address the internal use of data formats and data structure. Such general technology standards exist, but seem to be rarely used in the LO community.
5. It should be prescribed that the architecture of LOs be layered as a part of best practice, in order to separate data, presentation and application logics. This would enhance the level of decomposability and context independence, as is also pointed out in the paper by Pinkwart *et al.* (2005). Layering (or multi-tier architectures) is used frequently in many other areas of application/system development for the very same reasons.

6. Pedagogy should preferably be kept outside the LO in order to facilitate pedagogical context independence. It is suggested that the pedagogical model is added as LOs are assembled to form learning modules. Using such methodology, it becomes possible to do pedagogical contextualisation at a later stage in the authoring process, and enhance reusability of different components as well as components mutual pedagogical context independence.

According to Paulsson and Naeve (2006), in some cases there might be a need to add such 'instructional properties' inside LOs, but in such cases this should be handled in a separate layer, using standard specifications for that purpose, and not by hard coded implementations.

1.4. *Nationally Recognised LOs Quality Criteria*

This subsection is a short overview of the European nationally recognised (i.e., approved on the national level) LOs quality criteria. These LOs quality evaluation criteria (not only technological) are shortly presented below according to Kurilovas (2009).

Belgium (Flanders) LOs quality criteria are divided into two groups:

Scope criteria:

1. Aims and users (aims; target group).
2. Contents (curriculum area; relevant source; source type; granularity; no advertisements).
3. Geographical coverage (geographical limitations (availability – relevance); language).
4. Accessibility (technology; costs; registration).
5. Source data (metadata).

Quality criteria:

1. Content criteria (validity; accuracy; authority; actuality; substantial coverage; completeness).
2. Formal criteria (navigation; user support; use of ICT and standards).
3. Process criteria (integrity of information; site stability; platform stability).

Czech LOs quality criteria are as follows:

1. Content accuracy.
2. Connection with national curriculum.
3. Didactics/pedagogy accuracy.
4. Typographic rules.
5. Citations, copyright.

French LOs quality criteria are as follows:

1. Curriculum conforming.
2. Free rights for pedagogical use in the classroom and out of the classroom.
3. Respect of publishers engagements.
4. Provide access to expertise the resources.
5. Update information must be available.
6. Provide the possibility to block advertising (if included).

Lithuanian LOs quality criteria are as follows:
1. Methodical aspects.
2. User interface (incl. personalisation).
3. LOs arrangement possibilities.
4. Communication and collaboration possibilities and tools.
5. Technical features (incl. working stability).
6. Documentation.
7. Implementation and maintenance expenditure.

Norwegian LOs quality criteria are as follows:
1. Technical interoperability (file formats, standards; IMS QTI and CP, etc.).
2. Metadata tagging.
3. Accessibility (WCAG 1.0).
4. Parallel publishing (two forms of the Norwegian language exist).
5. Pedagogical issues.

Portugal LOs quality criteria are brought together under five different domains:
1. Technical.
2. Scientific.
3. Educational.
4. Language.
5. Values and attitudes.

Swedish LOs quality criteria are as follows:
1. Every web site must have source information and must not conflict with Swedish law.
2. The presentation must be clear and easy to navigate.
3. The information must be reliable: who is responsible for the information, why the information is published, how often it is updated, etc.
4. Other criteria deal with general interest, language and access requirements.

Swiss LOs quality criteria are divided into two groups:
National level:
1. Recommendation for description (application profile).
2. Technical tools.
3. Harvesting.

Local/partner level:
1. Selection and validation of resources according local criteria and culture.
2. Metadata check, enrichment and maintenance.
3. Publication (local).

1.5. *MELT Quality Criteria*

MELT (2008) set of the LOs quality criteria has been divided into five categories:
1. Pedagogical.
2. Usability.
3. Reusability.

4. Accessibility.
5. Production.

The list was by no means prescriptive and not all of the criteria can always be applied to all LOs. For example, some LOs may score strongly in terms of reusability because they include open source code that facilitates adaptation to different learning scenarios than the one originally intended. However, the same LOs might actually score poorly in terms of its interactivity. The set of the quality criteria, therefore, needed to be seen more as a minimum framework that should be used in a flexible way.

However, it is also important to appreciate that some very high-quality LOs may meet the specific needs of a national curriculum but may not always have the ability to be used as effectively (or maybe at all) by schools in other countries. For example, a text-heavy lesson plan in a minority European language may work splendidly in a national context but may simply be unusable by teachers in other countries.

With this in mind MELT has begun to develop quality criteria that are defined in terms of the extent to which learning content has the potential to 'travel well'; i.e., the extent to which LOs/assets can be easily used across national borders and in different curricula frameworks.

An initial assumption in MELT was that content is more likely to 'travel well' if it is:
- Modular: the parts of a content item are fully functional on their own.
- Adaptable: the LO can be modified, for instance from a configuration file, from a plain text file or because it is provided along with its source code or an authoring tool.

The cross-border reuse of content will be more likely if LOs:
- Have a strong visual element and users can broadly understand what is the intended learning objective or topic. E.g., LOs may have little or no text; and include animations and simulations that are self-explanatory or have just a few text labels or icons/buttons for start, stop, etc.
- Have been designed to be language customisable and are already offered in more than one language.
- Address curriculum topics that could be considered trans-national.
- Are adaptable from a technical (e.g., LOs are supplied along with an authoring environment or tools) or IPR perspective.

1.6. *Quality for Reuse Criteria*

A quality assurance strategy was implemented in Q4R (2008) scientific project initiated by Tele-University of Quebec to improve effectiveness, efficiency and flexibility of LOs as well as proper storing and retrieval strategies.

Q4R quality assurance strategies are organised into four main groups, namely organisational strategies, and then three strategies inspired by the life-cycle of a LO, that is from its conception to its use/reuse (adaptations) – before, during and after LO inclusion in the LO repository (LOR).

Strategies before LO inclusion in the LOR are based on the following principles:

- Only build or integrate LOs, which can be certified for quality.
- Interactive LOs are software and as such should answer to software quality criteria.

Strategies during LO inclusion in the LOR are based on:

- The principle of reducing form-filling.
- Use of guiding wizards, smart automatic and semi-automatic computer agents to assist in assuring technical interoperability.

Strategies after LO inclusion in the LOR are based on the following principles:

- Provide interesting and easily understood user statistics, such as stars, percentages, voting systems.
- Include recommendations for reuse by the user, both to the next user and the designer.

1.7. *European Learning Resource Exchange Service Validation in Lithuania*

FP6 CALIBRATE (CALIBRATE, 2008) project has been one of the most significant European scientific research projects aimed at the research and development of LOs interoperability and exchange. CALIBRATE portal's usability was evaluated in Lithuania in autumn 2007. This portal is the gates to European Learning Resource Exchange service for schools containing thousands LOs.

During the validation Lithuanian teachers were asked by the authors how they will use LOs from CALIBRATE portal. The results of the survey are as follows: most of the teachers have been using the found LO as illustration or include it in presentation (i.e., as learning asset; 88%), in project work (78%), as background information (76%), or include LO in test/task sheet (66%).

Project results have shown that the majority of the teachers prefer to reuse 'small' learning assets, and they intend to reuse the majority of assets in another way and in another learning context than was primarily designed by the developers.

CALIBRATE results have proven the authors' research results on LOs reusability as one of the main criteria for 'quality in use' evaluation of LOs (Kurilovas, 2009; 2007; Kurilovas and Dagiene, 2008; Dagiene and Kurilovas, 2008).

1.8. *The Comprehensive Set of Criteria for Learning Objects Technological Evaluation*

While analysing the aforementioned LOs evaluation criteria (Sections 1.2–1.7) it was necessary to exclude all evaluation criteria that do not deal directly with LOs technological quality problems on the one hand, and to estimate interconnected/overlapping criteria on the other.

This analysis has shown that all analysed sets of LOs evaluation criteria have a number of limitations from technological point of view:

- LORI, Paulsson and Naeve and MELT criteria do not examine different LO life cycle stages.
- Q4R set of criteria insufficiently examines technological evaluation criteria before LO inclusion in the LO repository.

- All these criteria insufficiently examine LO reusability (incl. 'Interoperability').

The nationally recognised Lithuanian set of LOs evaluation criteria (see Section 1.4) also has a number of limitations (Kurilovas, 2009, 2007; Kurilovas and Dagiene, 2008):

- All LOs and services (such as LORs, VLEs) are evaluated against the same criteria.
- No international standards compliant metadata related criteria are included.
- These criteria for e-content, e-activities and e-services do not reflect their reusability aspects overall.

It is obvious that a more comprehensive set of criteria for LOs technological evaluation is needed. This comprehensive set of criteria should include LO quality evaluation criteria suitable for different LO life cycle stages, including criteria before, during and after LO inclusion in the repository as well as LO reusability criteria (Kurilovas, 2009, 2007; Kurilovas and Dagiene, 2008). LO reusability criteria should have the same weight as the other criteria (Kurilovas, 2009).

Therefore the comprehensive set of criteria for LOs technological evaluation is proposed. This set of criteria is based on the flexible e-learning system approach (see Section 1.1) as well as on the analysis of LO quality evaluation criteria presented in Sections 1.2–1.7. It combines LORI, Paulsson and Naeve, MELT, Q4R and the other research results published in Kurilovas (2009, 2007), Dagiene and Kurilovas (2008), Kurilovas and Dagiene (2008).

This comprehensive set of criteria includes LOs technological evaluation criteria suitable for different LOs life cycle stages (before, during and after LO inclusion in the LOR), as well as LO reusability criteria. It combines both 'internal quality' criteria suitable for all LOs (such as 'Interoperability', 'Architecture', 'Working stability') and 'quality in use' criteria suitable for the particular projects or learners groups. Therefore this set of criteria is suitable for the expert evaluation of LOs 'quality in use' as well as the 'internal quality' (see Section 1.1). Teachers and learners are reputed as the main users groups here.

This original set of criteria is presented in Table 1.

The additional LO evaluation criteria interconnected with technological criteria could be as follows (Kurilovas, 2007):

- Licensing (clear rules, e.g., compliance with Creative Commons).
- Economic efficiency (taking into account probable LO reusability level).

2. Quality Evaluation Criteria of Virtual Learning Environments

The proposed comprehensive set of VLE evaluation criteria is based on the flexible personalised approach to creation of e-learning systems (Dagiene and Kurilovas, 2008, 2007; Kurilovas and Kubilinskienė, 2008a, 2008b) as well as mainly on two well-known VLE evaluation methods suitable for flexible personalised approach:

- Methodology of technical evaluation of learning management systems (LMSs) (Catalyst IT Ltd., 2004).
- Method of evaluation of open source e-learning platforms with the main focus is on adaptation issues (Graf and List, 2005).

Table 1

Technological quality evaluation criteria of learning objects (Kurilovas and Dagiene, 2008)

Criteria before LO inclusion in LOR	Narrow definition compliance		
	Reusability level: interoperability	Metadata accuracy Compliance with the main import/export standards (IMS)	
	Reusability level: decontextualisation	Is LO indivisible (atomic)? Does LO have a strong visual element?	LO aggregation (granularity) level Is LO modular?
	Reusability level: cultural and learning diversity principles	Is LO flexible (can be modified)? LO suitability for localisation LO internationalisation level	
	Reusability level: accessibility	Is LO designed for all? Compliance with accessibility standards (W3C)	
	LO architecture	Is LO architecture layered in order to separate data, presentation and application logics?	
	Working stability		
	Design and usability	Aesthetics Navigation User-friendly interface Information structurisation Personalisation	
Criteria during LO inclusion in LOR	Membership of contribution control strategies	Using LO harvesting Obligatory membership	
	Technical interoperability	Automatic verification of capability with known protocols Automatic metadata generation or simplified metadata tagging	
Criteria after LO inclusion in LOR	Retrieval quality	User should be able to retrieve LO in different ways	
	Information quality	Display information strategies Feedback techniques	

2.1. *General Evaluation Criteria of Virtual Learning Environments*

General evaluation criteria for VLEs were formulated in the 'Methodology of Technical Evaluation of LMSs' (or VLEs), which is a part of the Evaluation of Learning Management Software activity undertaken as part of the Open Source VLE project (Catalyst IT Ltd., 2004). The evaluation criteria expand on a subset of the criteria, focusing on the technological aspects of VLEs.

Overall VLEs technological evaluation criteria in conformity with this methodology are:

1. Overall architecture and implementation:
   - Scalability of the system.
   - System modularity and extensibility.
   - Possibility of multiple installations on a single platform.

- Reasonable performance optimisations.
- Look and feel is configurable.
- Security.
- Modular authentication.
- Robustness and stability.
- Installation, dependencies and portability.

2. Interoperability:
  - Integration is straightforward.
  - LMS/VLE standards support.
3. Cost of ownership.
4. Strength of the development community:
  - Installed base and longevity.
  - Documentation.
  - End-user community.
  - Developer community.
  - Open development process.
  - Commercial support community.
5. Licensing.
6. Internationalisation and localisation:
  - Localisable user interface.
  - Localisation to relevant languages.
  - Unicode text editing and storage.
  - Time zones and date localisation.
  - Alternative language support.
7. Accessibility:
  - Text-only navigation support.
  - Scalable fonts and graphics.
8. Document transformation.

The criteria 'Overall architecture and implementation', 'Interoperability', 'Internationalisation' and 'Accessibility' are suitable for the technological evaluation of VLEs (Kurilovas, 2005).

2.2. *Adaptation Evaluation Criteria of Virtual Learning Environments*

Graf and List (2005) have presented the open source e-learning platforms/VLEs evaluation method where the main focus has been on adaptation issues.

Before (Graf and List, 2005), adaptation received very little coverage in VLEs. The extended VLE is utilised in an operational teaching environment. Therefore, the overall functionality of VLE is as important as the adaptation capabilities. Graf and List (2005) evaluation treats both issues.

Graf and List (2005) evaluation is also based on the qualitative weight and sum approach. After a pre-evaluation phase, nine platforms/VLEs have been analysed in detail.

The detailed evaluation approach has been focused on the adaptation category and its results.

Adaptation Capabilities section in Graf and List (2005) has been concentrated attention on:

- Adaptability.
- Personalisation.
- Extensibility.
- Adaptivity capabilities of the platforms.

However the study of Graf and List (2005) has been focused on customisable adaptation only, which can be done without programming skills:

- Adaptability includes all facilities to customise the platform for the educational institution's needs (e.g., the language or the design).
- Personalisation aspects indicate the facilities of each individual user to customise his/her own view of the platform.
- Extensibility is, in principle, possible for all open source products. Nevertheless, there can be big differences. For example, a good programming style or the availability of a documented application programming interfaces (API) are helpful.
- Adaptivity capabilities of the platforms indicate all kinds of automatic adaptation to the individual user's needs (e.g., personal annotations of LOs or automatically adapted content; Graf and List, 2005).

2.3. *The Comprehensive Set of Criteria for Technological Evaluation of Virtual Learning Environments*

While analysing the aforementioned sets of criteria for VLEs evaluation (Sections 2.1–2.2) it was necessary to exclude all evaluation criteria that do not deal directly with VLEs technological quality problems on the one hand, and to estimate interconnected/overlapping criteria on the other.

This analysis has shown that the both analysed VLE technological evaluation methods have a number of limitations:

- The method developed in Catalyst IT Ltd. (2004) practically does not examine adaptation capabilities criteria.
- The method proposed by Graf and List (2005) insufficiently examines general technological criteria.

Therefore a more comprehensive set of criteria for VLE technological evaluation is needed.

It should include general technological evaluation criteria based on modular approach and interoperability, as well as adaptation capabilities criteria. VLE adaptation capabilities criteria should have the same weight as the other criteria (Kurilovas and Dagiene, 2008).

Therefore the original comprehensive set of criteria for VLEs technological evaluation is proposed. This comprehensive set of criteria is presented in Table 2. It combines the general technological evaluation criteria based on the modular approach and inter-

Table 2

Virtual Learning Environments technological quality evaluation criteria (Kurilovas and Dagiene, 2008)

General criteria	1. Overall architecture and implementation	Scalability Modularity (of the architecture) Possibility of multiple installations on a single platform Reasonable performance optimisations Look and feel is configurable Security Modular authentication Robustness and stability Installation, dependencies and portability
	2. Interoperability	Integration is straightforward VLE standard support
	3. Internationalisation and localisation	Localisable user interface Localisation to relevant languages Unicode text editing and storage Time zones and date localisation Alternative language support
	4. Accessibility	Text only navigation support Scalable fonts and graphics
Adaptation criteria	5. Adaptability (facilities to customise for the educational institution's needs)	Language Design
	6. Personalisation aspects (facilities of each individual user to his/her own view of the platform)	
	7. Extensibility	Good programming style Availability of a documented API
	8. Adaptivity (all kinds of automatic adaptation to the individual user's needs)	Personal annotations of LOs Automatically adapted content

operability (Kurilovas, 2005), as well as adaptation capabilities criteria (Kurilovas and Dagiene, 2008).

This set of criteria is suitable for the expert evaluation of both VLEs 'quality in use' criteria (see criteria 1–4) and 'internal quality' (see criteria 5–8). Teachers and learners are reputed as the main users groups here.

3. Multiple Criteria Evaluation of Virtual Learning Environments

3.1. *Ratings of Quality Evaluation Criteria*

Scientists who have explored quality of software consider that there exists no simple way to evaluate functionality characteristics of internal quality of software.

According to Gasperovic and Caplinskas (2006), it is a hard and complicated task, which requires relatively high time and labour overheads.

According to Zavadskas and Turskis (2008), "each alternative in multi-criteria decision making problem can be described by a set of criteria. Criteria can be qualitative and quantitative. They usually have different units of measurement and different optimisation direction".

The comprehensive sets of criteria suitable for multiple criteria expert evaluation (decision making) of LOs (see Table 1) and VLEs (see Table 2) are proposed in the present article.

Evaluation criteria ratings (values) are also necessary for the evaluation methods suitable for the practical expert evaluation of LOs and VLEs.

It is proposed to the use the universal set of the ratings (values) of the quality evaluation criteria (Wyles *et al.*, 2006) for the practical evaluation of the all main e-learning system components: LOs, their repositories and VLEs.

This set of ratings is based on the analysis of the level of criterion/feature support and the level of modification (e.g., time, money) needed to reach the desired level of support (Table 3).

This set of ratings (values) is clearer in comparison with the ratings developed by (Catalyst IT Ltd., 2004) and more convenient in comparison with the ratings proposed by Graf and List (2005), Kurilovas and Dagiene (2008).

Each criterion is proposed to be given a rating (value) to be used when evaluating LOs, repositories and VLEs.

Major criteria have to be broken down into sub-criteria with each sub-criterion also having a rating.

The rating range is 0–4, with 0 being the lowest and 4 being of the highest value.

3.2. *Experimental Evaluation of Virtual Learning Environments for the General Case*

The VLEs quality evaluation criteria presented in Table 2 and the ratings (values) of these criteria presented in Table 3 are used here to practically evaluate three most popular open source VLEs (ATutor, Ilias and Moodle) against the technological (both General and Adaptation) criteria.

Table 3
Ratings (values) of quality evaluation criteria (Kurilovas and Dagiene, 2008)

Rating	Explanation
0	Failed or feature does not exist
1	Has poor support and/or it can be done but with significant effort
2	Fair support but needs modification to reach the desired level of support
3	Good support and needs a minimal amount of effort
4	Excellent support and meets the criteria off the scale, minimal effort

In this case we invoke the generic principle that all the weights of evaluation criteria are equal 1.

Therefore we can use the expert's (decision maker's) additive utility function

$$f(X) = \sum_{i=1}^{m} f_i(X), \tag{1}$$

where

$$f_i(X_j) = \{0, 1, \ldots, 4\}$$

is the rating (see Table 3) of the criterion i for the each of the 3 examined alternatives X_j:

X_1 – ATutor, X_2 – Ilias, and X_3 – Moodle.

Here i are the order numbers of the VLEs quality evaluation criteria presented in Table 2. First four of these criteria are general VLE quality criteria, and the other four – VLE adaptation criteria.

The major is the meaning of the utility function (1) the better is the quality of the VLE.

All the alternatives were evaluated against the criteria presented in Table 2 using the criteria ratings presented in Table 3.

The three VLEs technological evaluation summary is presented in Table 4.

In conformity with this experimental evaluation results, Moodle is the best VLE from technological point of view in general case. This alternative has shown the highest ratings of both 'internal quality' evaluation (see "General criteria ratings") and 'quality in use' evaluation (see "Adaptation criteria ratings").

Table 4

VLE technological evaluation summary (Kurilovas and Dagiene, 2008)

Quality evaluation criteria	ATutor	Ilias	Moodle
General ('internal quality') criteria ratings			
1. Architecture and implementation	2	1	4
2. Interoperability	3	3	2
3. Internationalisation and localisation	1	2	3
4. Accessibility	4	1	2
General evaluation rating:	*10*	*7*	*11*
Adaptation ('quality in use') criteria ratings			
5. Adaptability	1	2	3
6. Personalisation	3	3	2
7. Extensibility	3	4	4
8. Adaptivity	1	0	1
Adaptation evaluation rating:	*8*	*9*	*10*
Total evaluation rating $f(X)$:	*18*	*16*	*21*

The main Moodle strengths are its simple and complete modular architecture, excellent scalability, good security, modular authentication, robustness and stability, excellent portability and localisation possibilities. The adaptation strengths of Moodle are realisation of communication tools, creation and administration of LOs, comprehensive didactical concepts, tracking of data, and outstanding usability.

3.3. *Experimental Evaluation of Virtual Learning Environments for the Particular Learner Needs*

Overall, if we want evaluate (or to optimise) LOs and VLEs (or the overall e-learning system) for the particular learner needs, besides technological evaluation criteria we have to use an extensive additional set of pedagogical quality criteria of LOs (see Section 1) and VLEs (e.g., based on Conversational Framework presented in Kurilovas, 2005) as well as VLEs organisational criteria (e.g., based on Viable System Model presented in Kurilovas, 2005).

In particular, if we want to evaluate (or optimise) the technological quality of VLE (see Table 2) for the particular learner (i.e., to personalise his/her learning process in the best way in conformity with their prerequisites, preferred learning speed and methods etc.), we should use the general VLE expert evaluation method presented by the function (1) together with the weights of evaluation criteria. The weight of the evaluation criterion reflects the experts' opinion on the criterion's importance level in comparison with the other criteria for the concrete learner/user.

In this case we have the multiple criteria optimisation task using the weights of the alternatives' evaluation criteria.

There are some suitable methods exist to solve such kind of tasks.

E.g., Dzemyda and Petkus (2001) have proposed the method of application of computer network to solve the complex applied multiple criteria optimisation problems. Zavadskas and Turskis (2008) have proposed logarithmic normalisation method in games theory. Dzemyda and Šaltenis (1994) have proposed multiple criteria decision support system. Three interactive methods of increasing complexity were realised in this system (Dzemyda and Šaltenis, 1994):

- Paired comparisons of alternatives.
- Pareto.
- Fuzzy.

Let us explore the task of multiple criteria optimisation of VLEs for the particular project or learners group using VLEs quality criteria ratings and their weights. Let us consider the following weights

$$a_i = \{1, \ldots, 4\}$$

for the quality evaluation criteria

$$i = \{1, \ldots, 8\}$$

(see Table 3).

A possible decision could be to transform multi-criteria task into one-criterion task obtained by adding all criteria together with their weights. In this case we have more complex experts' additive utility function:

$$f(X) = \sum_{i=1}^{m} a_i f_i(X), \quad a_i > 0. \tag{2}$$

The major is the meaning of the utility function (2) the better VLE meets the concrete learner needs. For example, if the experts (decision makers) want to select the most suitable VLE for the students with special education needs/disabilities, they should choose higher weights for the particular criteria such as e.g., Accessibility (e.g., measuring weight $a_4 = 3$) and Personalisation (e.g., measuring weight $a_6 = 2$). In this case all the other criteria weights are measured $a_i = 1$.

In this case the experts should find that differently from the general case (i.e., without using the weights of the criteria reflecting the particular learners' preferences) ATutor is the optimal VLE for the students with special needs.

In conformity with the function (2) ATutor's General evaluation rating will be 18, and Adaptation evaluation rating will be 11. Total ATutor evaluation rating will be $f(X_1) = 29$, while $f(X_2) = 21$ (for Ilias) and $f(X_3) = 27$ (for Moodle).

Such approach has never been applied for solving of the e-learning systems and components optimisation tasks before.

4. Conclusions

The proposed comprehensive learning objects' and virtual learning environments' technological evaluation methods combine both 'internal quality' criteria that describe the quality of software independently from any particular context of its use, and 'quality in use' evaluation criteria suitable for the particular project or learner. Therefore these methods are more suitable for the multiple criteria expert evaluation of e-learning software quality in comparison with the other e-learning software quality evaluation methods.

VLEs quality evaluation method (i.e., the experts' additive utility function (2)) including VLEs comprehensive technological evaluation criteria together with their ratings (values) and the experts preferred weights is suitable to apply for the practical evaluation of the VLEs in order to meet the particular learner needs.

The research has shown that in the general case Moodle is the best VLE from technological point of view, but ATutor is the best VLE for the special needs students.

References

CALIBRATE (2008). FP6 IST CALIBRATE (Calibrating eLearning in Schools) project web site. Available: `http://calibrate.eun.org`.

Catalyst IT Ltd. (2004). *Technical Evaluation of Selected Learning Management Systems.* Available: `https://eduforge.org/docman/view.php/7/18/LMS%20Technical%20Evaluation%20-%20May04.pdf`.

Dagienė, V., Kurilovas, E. (2007). Design of Lithuanian digital library of educational resources and services: the problem of interoperability. *Information Technologies and Control*, 36(4), 402–411.

Dagienė, V., Kurilovas, E. (2008). *Information Technologies in Education*: *Experience and Analysis*. Institute of Mathematics and Informatics, Vilnius (in Lithuanian).

Dzemyda, G., Petkus, T. (2001). Application of computer network to solve the complex applied multiple criteria optimisation problems. *Informatica*, 12(1), 45–60.

Dzemyda, G., Šaltenis, V. (1994). Multiple criteria decision support system: methods, user's interface and applications. *Informatica*, 5(1–2), 31–42.

Gasperovic, J., Caplinskas, A. (2006). Methodology to evaluate the functionality of specification languages. *Informatica*, 17(3), 325–346.

Graf, S., List, B. (2005). An evaluation of open source e-learning platforms stressing adaptation issues. Presented at: *ICALT 2005*.

Haughey, M., Muirhead, B. (2005). Evaluating learning objects for schools. Available:
`http://www.usq.edu.au/electpub/e-jist/docs/vol8_no1/fullpapers/Haughey_`
`Muirhead.pdf.`

Institute of Mathematics and Informatics (2005). Research on computer learning tools and virtual learning environments implementation in vocational education. Scientific research report.
Available: `http://www.emokykla.lt/lt.php/tyrimai/194.`

ISO/IEC 14598-1:1999 (1999). *Information Technology – Software Product Evaluation – Part 1*: *General Overview*, 1st ed., 1999-04-15.

Kurilovas, E. (2005). Several aspects of technical and pedagogical evaluation of virtual learning environments. *Informatics in Education*, 4(2), 215–252.

Kurilovas, E. (2007). Digital library of educational resources and services: evaluation of components. *Informacijos mokslai*, 42–43, 69–77.

Kurilovas, E. (2009). Learning objects reusability and their adaptation for blended learning. In: *Proceedings of the 5th International Conference on Networking and Services* (*ICNS 2009*). Valencia, Spain, pp. 542–547.

Kurilovas, E., Dagiene, V. (2008). Learning objects and virtual learning environments technical evaluation tools. In: *Proceedings of the 7th European Conference on E-Learning* (*ECEL'08*), Vol. 2. Agia Napa, Cyprus, pp. 24–33.

Kurilovas, E., Kubilinskienė, S. (2008a). Interoperability framework for components of digital library of educational resources and services. *Informacijos mokslai*, 44, 89–98.

Kurilovas, E., Kubilinskiene, S. (2008b). Analysis of Lithuanian LOM repository strategies, standards and interoperability. In: *Proceedings of the 2nd International Workshop on Search and Exchange of E-Learning Materials* (*SE@M'08*) *within the 3rd European Conference on Technology Enhanced Learning* (*EC–TEL'08*), Vol. 385. Maastricht, Netherlands,
URL: `http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-`
`385/.`

McGreal, R. (2004). Learning objects: a practical definition. *International Journal of Instructional Technology and Distance Learning*, 1, 21–32.

MELT (2008). *EU eContentplus Programme's MELT* (Metadata Ecology for Learning and Teaching) project web site. Available: `http://melt-project.eun.org.`

Paulsson, F., Naeve, A. (2006). *Establishing Technical Quality Criteria for Learning Objects*.
Available: `http://www.frepa.org/wp/wp-content/files/Paulsson-Establ-Tech-`
`Qual_finalv1.pdf.`

Pinkwart, N., Malzahn, N., Westheide, D., Hoppe, U. (2005). Community support based on thematic objects and similarity search. In: *12th International Conference on Artificial Intelligence in Education*. Amsterdam, Holland.

Q4R (2008). Quality for Reuse (Q4R) project web site. Available: `http://www.q4r.org.`

Van Assche, F., Vuorikari, R. (2006). A framework for quality of learning resources. In: *European Handbook for Quality*. CEDEFOP.

Vargo, J., Nesbit, J.C., Belfer, K., Archambault, A. (2003). Learning object evaluation: Computer mediated collaboration and inter–rater reliability. *International Journal of Computers and Applications*, 25(3), 198–205.

Wiley, D.A. (2000). *Connecting Learning Objects to Instructional Design Theory: A Definition, a Metaphor, and a Taxonomy*. Utah State University. Available: `http://www.reusability.org/read/`.

Wyles, R., *et al.* (2006). *Technical Evaluation of Selected Open Source Repository Solutions*. Available: `http://eduforge.org/docman/view.php/131/1062/Repository%20Evaluation%20Document.pdf`.

Zavadskas, E.K., Turskis, Z. (2008). A new logarithmic normalization method in games theory. *Informatica*, 19(2), 303–314.

E. Kurilovas is a doctor of technological sciences (informatics engineering). He is a research scientist in the Institute of Mathematics and Informatics and the head of International Training Department of the Centre for IT in Education of the Ministry of Education and Science of Lithuania, and an associate professor in Vilnius Gediminas Technical University.

His main scientific interests are: metadata, interoperability and standards; technology-enhanced learning problems analysis; e-learning systems architecture; learning objects, their repositories and virtual learning environments usability; e-learning systems' components quality evaluation criteria; optimisation of learning process.

V. Dagienė is head of Department of Informatics Methodology at the Institute of Mathematics and Informatics as well as professor at Vilnius University. She has published over 120 scientific papers and the same number of methodical works, has written more than 50 textbooks in the field of informatics and ICT for primary and secondary education. Her main research topics are informatics didactics, especially teaching algorithms and programming, technology enhanced learning, computing education research, localization of software.

She is national representative of the Technical Committee of IFIP for Education (TC3), member of the Group for Informatics in Secondary Education (WG 3.1) and for Research (WG 3.3) of IFIP, member of the European Logo Scientific Committee, member of International Committee of Olympiads in Informatics. She is Executive Editor of international journals "Informatics in Education" and "Olympiads in Informatics".

El. mokymosi sistemų ir komponentų daugiakriterinis lyginamasis vertinimas

Eugenijus KURILOVAS, Valentina DAGIENĖ

Straipsnyje nagrinėjamos elektroninio mokymosi sistemų komponentų (t.y. mokymosi objektų (MO) ir virtualiųjų mokymosi aplinkų (VMA)) daugiakriterinio kokybės vertinimo problemos. Straipsnio tikslas – išanalizuoti MO ir VMA kokybės vertinimo metodus ir sukurti išsamesnius ir labiau visapusiškus metodus, pagrįstus mokymosi individualizavimo aspektais. Straipsnyje MO ir VMA kokybės vertinimo kriterijai laikomi optimizavimo parametrais. Yra analizuojamos kelių optimizavimo metodų taikymo galimybės. Išsamiausiai remiamasi naudingumo funkcija naudojant vertinimo kriterijų įverčius ir svorius. Šiais naujais aspektais straipsnis skiriasi nuo kitų šios srities mokslinių darbų.