Australasian Journal of
Educational Technology

2008, 24(5), 574-591

A multi-component model for assessing learning objects: The learning object evaluation metric (LOEM)

Robin H. Kay and Liesel Knaack
University of Ontario Institute of Technology

While discussion of the criteria needed to assess learning objects has been extensive, a formal, systematic model for evaluation has yet to be thoroughly tested. The purpose of the following study was to develop and assess a multi-component model for evaluating learning objects. The Learning Object Evaluation Metric (LOEM) was developed from a detailed list of criteria gathered from a comprehensive review of the literature. A sample of 1113 middle and secondary students, 33 teachers, and 44 learning objects was used to test this model. A principal components analysis revealed four distinct constructs: interactivity, design, engagement, and usability. These four constructs showed acceptable internal and inter-rater reliability. They also correlated significantly with student and teacher perceptions of learning, quality, and engagement. Finally, all four constructs were significantly and positively correlated with student learning performance. It is reasonable to conclude that the LOEM is reliable, valid, and effective approach to evaluating the effectiveness of learning objects in middle and secondary schools.

Overview

Considerable discussion and debate has been directed toward establishing an acceptable definition of learning objects (Agostinho et al., 2004; Butson, 2003; Friesen, 2001; Gibbons, Nelson & Richards, 2002; McGreal, 2004; Parrish, 2004; Wiley et al., 2004), however, consensus has yet to be reached. Multiple definitions and the variety of potential learning objects available have made it challenging for researchers to create a systematic tool for evaluating quality and effectiveness (Haughey & Muirhead, 2005; Nurmi & Jaakkola, 2005, 2006a, 2006b; Sosteric & Hesemeirer, 2004). Nonetheless, it is essential that a reliable, valid assessment model is developed for at least three reasons (Bradley & Boyle, 2004; Downes, 2003).

First, it is unlikely that educators will use learning objects extensively without some assurance of value and quality (Vargo, Nesbit, Belfer & Archambault, 2002). Second, one of the main premises for using learning objects, namely reuse, is compromised without some sort of evaluation metric (Downes, 2003; Malcolm, 2005). Third, an effective assessment tool could greatly reduce search time for users who would only need to examine highly rated learning objects (Koppi, Bogle & Bogle, 2004). Ultimately, the foremost evaluation question that needs to be addressed is "what key features of a learning object support and enhance learning?" (Sosteric & Hesemeirer, 2002). The purpose of the following study, then, is to develop and assess a multi-component model for evaluating learning objects.

Definition of learning objects

While consensus about the definition of learning objects has yet to be reached, it is important to establish an acceptable working definition in order to move forward in developing an evaluation metric. Original definitions focused on technological issues such accessibility, adaptability, the effective use of metadata, reusability, and standardisation (e.g. Downes, 2003; Koppi, Bogle & Bogle, 2005; Muzio, Heins & Mundell, 2002; Nurmi & Jaakola, 2006b; Parrish, 2004; Siqueira, Melo & Braz, 2004). More recently, researchers have emphasised learning qualities such as interaction and the degree to which the learner actively constructs knowledge (Baruque & Melo, 2004; Bennett & McGee, 2005; Bradley & Boyle, 2004; Caws, Friesen & Beaudoin, 2006; Cochrane, 2005; McGreal, 2004; Kay & Knaack, in press; Sosteric & Hesemeirer, 2002; Wiley et al., 2004).

While both technical and learning based definitions offer important qualities that can contribute to the success of learning objects, evaluation tools focusing on learning are noticeably absent (Kay & Knaack, in press). In order to address a clear gap in the literature on evaluating learning objects, a pedagogically focused definition has been adopted for the current study based on a composite of previous definitions. Key factors emphasised included interactivity, accessibility, a specific conceptual focus, meaningful scaffolding, and learning. Learning objects, then, are operationally defined as "interactive web-based tools that support the learning of specific concepts by enhancing, amplifying, and/or guiding the cognitive processes of learners". To view specific examples of learning objects used by teachers in this study, see Appendix C at Kay & Knaack (2008c).

Previous approaches to evaluating learning objects

Theorists and researchers have advocated a number of approaches for evaluating learning objects including an emphasis on reuse (e.g. Convertini et al., 2005; Cochrane, 2005; Del Moral & Cernea, 2005), standards (e.g. McGreal et al., 2004; Nesbit & Belfer, 2004; Williams, 2000), converging opinions of various stakeholders (e.g. Nesbit & Belfer, 2004; Vargo et al., 2003), design (e.g. Bradley & Boyle, 2004; Krauss & Ally, 2005; Maslowski & Visscher, 1999), use (e.g. Bradley & Boyle, 2004; Buzetto-More & Pinhey, 2006, Kenny et al., 1999), and learning outcomes (e.g. Adams, Lubega, & Walmsley, 2005; Bradley & Boyle, 2004; MacDonald et al., 2005). One fundamental problem with these proposed models is that they are unsupported by empirical evidence (e.g. Buzetto-More & Pinhey, 2006; Cochrane, 2005; Krauss & Ally, 2005).

While the vast majority of learning object evaluation has been informal (Adams et al., 2004; Bradley & Boyle, 2004; Clarke & Bowe, 2006a, 2006b; Concannon et al., 2005; Fournier-Viger et al., 2006; Howard-Rose & Harrigan, 2003; Kenny et al., 1999; Lopez-Morteo & Lopez, 2007; MacDonald et al., 2005), several researchers have discussed and analysed comprehensive models for evaluating learning objects (Cochrane, 2005; Haughey & Muirhead, 2005; Kay & Knaack, 2005; Krauss & Ally, 2005; Nesbit & Belfer, 2004).

Haughey & Muirhead (2005) looked at a model for assessing learning objects which included the following criteria: integrity/accuracy of material, clarity of instructions, ease of use, engagement, scaffolding, feedback, help, visual/auditory, clarity of learning objectives, identification of target learners, prerequisite knowledge, appropriateness for culture and ability to run independently. While comprehensive, this framework has never been tested.

Nesbit and Belfer (2004) refer to the learning object review instrument (LORI) which includes nine items: content quality, learning goal alignment, feedback and adaptations, motivation, presentation design (auditory and visual), interaction (ease of use), accessibility (learners with disabilities), reusability, and standards. This instrument has been tested on a limited basis (Krauss & Ally, 2005; Vargo et al., 2003) for a higher education population, but the impact of specific criteria on learning has not been examined.

One of the better known evaluation models, developed by MERLOT, focuses on quality of content, potential effectiveness as a teaching-learning tool, and ease of use. Howard-Rose & Harrigan (2003) tested the MERLOT model with 197 students from 10 different universities. The results were descriptive and did not distinguish the relative impact of individual model components. Cochrane (2005) tested a modified version of the MERLOT evaluation tool that looked at reusability, quality of interactivity, and potential for teaching, but only final scores were tallied, so the impact of separate components could not be determined. Finally, the reliability and validity of the MERLOT assessment tool has yet to be established.

Kay and Knaack (2005, 2007a) developed an evaluation tool based on a detailed review of research on instructional design. Specific assessment categories included organisation/ layout, learner control over interface, animation, graphics, audio, clear instructions, help features, interactivity, incorrect content/ errors, difficulty/ challenge, useful/ informative, assessment and theme/ motivation. The evaluation criteria were tested on a large secondary school population. Reliability and validity were determined to be acceptable and the impact of individual features was able to be assessed. Students benefited more if they were comfortable with computers, the learning object had a well organised layout, the instructions were clear, and the theme was fun or motivating. Students appreciated the motivational, interactive, and visual qualities of learning objects most.

In summary, while most existing models of learning object evaluation include a relatively comprehensive set of evaluation criteria, with the exception of Kay & Knaack (2005, 2007a, 2007b, in press), the impact of individual features is not assessed and reliability and validity estimates are not provided. Proposed models, then, are largely theoretical at this stage in the evolution of learning object assessment.

Learning object evaluation metric (LOEM)

The model for evaluating learning objects in this study is based on a comprehensive review of the literature on instructional design (see Table 1) and key learning object evaluation models used previously (Cochrane, 2005; Haughey & Muirhead, 2005; Howard-Rose & Harrigan, 2003; Kay and Knaack, 2005, 2007a; Nesbit and Belfer, 2004). After considerable discussion and evaluation of the literature by three external experts in the area of learning objects, five main criteria were identified and include interactivity, design, engagement, usability, and content. With respect to *interactivity*, key components considered include promoting constructive activity, providing a user with sufficient control, and level of interactivity. The underlying theme is that learning objects should provide rich activities that open up opportunities for action, rather than prescribed pathways of learning (Brown & Voltz, 2005). When looking at *design*, investigators have focussed on layout, degree of personalisation, quality of graphics, and emphasis of key concepts. Evaluation of *engagement* has incorporated difficulty level, theme, aesthetic appeal, feedback, and inclusion of multimedia. *Usability* involves overall ease of use, clear instructions, and navigation. Finally, with *content*, the predominant features looked at are the integrity and accuracy of the material presented. Detailed references for each of the five main evaluation criteria observed are given in Table 1.

Table 1: Criteria for evaluating learning objects

Main criteria	Sub category	References
Interactivity	Constructive activity	Akpinar & Bal (2006); Baser (2006); Gadanidis et al. (2004); Jaakkola & Nurmi (2004); Jonassen (2006); Ohl (2001); van Marrienboer & Ayres (2005)
	Control	Deaudelin et al. (2003); Koohang & Du Plessis (2004); Nielson (2003); Ohl (2001)
	Level of interactivity	Cochrane (2005); Convertini et al. (2005); Lim et al. (2006); Lin & Gregor (2006); Metros (2005); Ohl (2001); Oliver & McLoughlin (1999); van Marrienboer & Ayres (2005)
Design	Layout	Buzetto-More & Pinhey (2006); Del Moral & Cernea (2005); Kay & Knaack (2005)
	Personalisation	Deaudelin et al. (2003)
	Quality of graphics	Koohang & Du Plessis (2004); Lin & Gregor (2006)
	Emphasis of key concepts	Gadanidis et al. (2004)
Engagement	Difficulty level	Haughey & Muirhead (2005)
	Theme	Brown & Voltz (2005); Haughey & Muirhead (2005); Jonassen (2006); Kay & Knaack (2005); Lin & Gregor (2006); Macdonald et al. (2005); Reimer & Moyer (2005); Van Zele et al. (2003)
	Aesthetics	Koohang & Du Plessis (2004);
	Feedback	Brown & Voltz (2005); Buzetto-More & Pinhey (2006); Haughey & Muirhead (2005); Koohang & Du Plessis (2004); Nesbit & Belfer (2004); Nielson (2003); Reimer & Moyer (2005)
	Multimedia	Brown & Voltz (2005); Gadanidis et al. (2004); Haughey & Muirhead (2005); Nesbit & Belfer (2004); Oliver & McLoughlin (1999)
Usability	Overall ease of use	Haughey & Muirhead (2005);Koohang & Du Plessis (2004); Lin & Gregor (2006); Macdonald et al. (2005); Nesbit & Belfer (2004); Schell & Burns (2002); Schoner et al. (2005)
	Clear instructions	Haughey & Muirhead (2005); Kay & Knaack (2005); Nielson (2003)
	Navigation	Concannon et al. (2005); Koohang & Du Plessis (2004); Lim et al. (2006)
Content	Accuracy	Haughey & Muirhead (2005); Macdonald et al. (2005)
	Quality	Nesbit & Belfer (2004); Schell & Burns (2002)

Purpose of this study

An elaborate, detailed set of criteria for evaluating learning objects has been discussed by theorists (see Table 1), but the impact of these criteria on learning behaviour has yet to be tested. empirically The purpose of this study was to assess systematically a multi-component model for evaluating learning objects.

Method

Overview

A variety of methods have been employed to evaluate learning objects, including focus groups (e.g. Krauss & Ally, 2005), informal observation (e.g. Clarke & Bowe, 2006a; 2006b; Fournier-Viger et al., 2006; MacDonald et al., 2005), general description (McCormick & Li, 2005; Vargo et al., 2003), interviews (e.g. Bradley & Boyle, 2004; Kenny et al., 1999), surveys, and think aloud protocols. In several cases, multiple tools have been used (e.g. Bradley & Boyle, 2004; Kenny et al., 1999; Van Zele, et al., 2003). However, a comprehensive review of 24 articles reveals a number of methodological problems that need to be addressed.

First, the majority of researchers approach data collection and analysis in an informal, somewhat *ad hoc* manner, making it challenging to generalise results observed (e.g. Clarke & Bowe, 2006a; 2006b; Fournier-Viger et al., 2006; MacDonald et al., 2005). Second, only two studies used some kind of formal statistical analysis to evaluate learning objects (Kay & Knaack, 2005; Van Zele et al., 2003). While qualitative research is valuable, it is important to include quantitative methodology, if only to establish triangulation. Third, reliability and validity estimates are rarely presented, thereby reducing confidence in any conclusions made (e.g. Bradley & Boyle, 2004; Clarke & Bowe, 2006a; 2006b; Fournier-Viger et al., 2006; Kenny et al., 1999; MacDonald et al., 2005; McCormick & Li, 2005; Vargo et al., 2003). Fourth, the sample size is often small and poorly described (e.g. Adams, et al., 2004; Bradley & Boyle, 2004; Cochrane, 2005; Kenny et al., 1999; Krauss & Ally, 2005; MacDonald et al., 2005; Van Zele et al., 2003). Finally, most research has focussed on a single learning object (e.g. Anderson, 2003; Van Zele et al., 2003; Vargo et al., 2003). It is critical, though, to test any evaluation tool on a wide range of learning objects.

In order to assure the quality and confidence of results reported, the following steps were taken in the current study:

1. a large, diverse, sample was used;
2. a wide range of learning objects was tested;
3. the design of the evaluation tools was based on a thorough review and categorisation of the learning object literature and instructional design research;
4. reliability and validity estimates were calculated;
5. formal statistics were used where applicable; and
6. a measure of learning performance was used.

Sample

Students
The sample consisted of 1113 students (588 males, 525 females), 10 to 22 years of age (M = 15.5, SD = 2.1), from both middle (n=263) and secondary schools (n=850). The population base spanned three different boards of education, six middle schools, 15 secondary schools, and 64 different classrooms. The students were selected through convenience sampling and had to obtain signed parental permission to participate.

Teachers
The sample consisted of 33 teachers (12 males, 21 females) and 64 classrooms (a number of teachers used learning objects more than once). These teachers had 0.5 to 33

years of teaching experience (*M* = 9.0, *SD* = 8.2) and came from both middle (n=6) and secondary schools (n=27). Most teachers taught mathematics (n=16) or science (n=15). A majority of the teachers rated their ability to use computers as strong or very strong (n=25) and their attitude toward using computers as positive or very positive (n=29), although, only six teachers used computers in their classrooms more than once a month.

Learning objects
In order to simulate a real classroom as much as possible, teachers were allowed to select any learning object they thought was appropriate for their curriculum. As a starting point, they were introduced to a wide range of learning objects located at the LORDEC website (LORDEC, 2008b). Sixty percent of the teachers selected learning objects from the LORDEC repository - the remaining teachers reported that they used *Google*. A total of 44 unique learning objects were selected covering concepts in biology, Canadian history, chemistry, general science, geography, mathematics, and physics (see Appendix C at Kay and Knaack (2008c) for the full list).

Procedure

Data collection
Teachers from three boards of education volunteered to use learning objects in their classrooms. Each teacher received a half day of training in November on how to choose, use, and assess learning objects (see LORDEC (2008b) for more details on the training provided). They were then asked to use at least one learning object in their classrooms by April of the following year. Email support was available throughout the duration of the study. All students in a given teacher's class used the learning object that the teacher selected, but only those students with signed parental permission forms were permitted to fill in an anonymous, online survey about their use of the learning object. In addition, students completed a pre- and post-test based on the content of the learning object.

Scale item analysis
Four teachers were trained over 3 half-day sessions on using the Learning Object Evaluation Metric (LOEM) (see Appendix B at Kay and Knaack (2008b) for details) to assess 44 learning objects. In session one (5 hours), two instructors and the four teacher raters discussed and used each item in the LOEM to assess a single learning object (3 hours). A second learning object was evaluated and discussed by the group (1 hour). The four teachers were then instructed to independently rate four more learning objects at home over the following two days.

The group then met a second time to discuss the evaluations completed at home (4 hours). Teachers were asked to re-assess all previously assessed learning objects based on the conclusions and adjustments agreed upon in the discussion. They were also asked to rate 10 more learning objects.

Three days later, the group met for a final time to discuss the evaluation of three more learning objects, chosen at random (4 hours). All teacher raters felt confident in evaluating the remaining learning objects and completed the 44 evaluations within the next six to seven days. Inter-rater reliability estimates (within one point) were as follows: rater 1 and rater 2, 96%; rater 1 and rater 3, 94%; rater 1 and rater 4, 95%; rater 2 and rater 3, 95%; rater 2 and rater 4, 96%; and rater 3 and rater 4, 95%.

Context in which learning objects were used
The mean amount of time spent on the learning object component of the lesson was 35.4 minutes (*SD* = 27.9, ± 6.8 minutes) with a range of 6 to 75 minutes. The most frequent reasons that teachers chose to use learning objects were to review a previous concept (n=34, 53%), to provide another way of looking at a concept (n=32, 50%), motivate students (n=28, 44%), and to introduce or explore a new concept before a lesson (n=20, 31%). Teachers rarely chose to use learning objects to teach a new concept (n=9, 14%), explore a new concept after a lesson (n=4, 6%), or to extend a concept (n=1, 2%).

Almost all teachers (n=59, 92%) chose to have students work independently on their own computers. With respect to introducing the learning object, 61% (n=39) provided a brief introduction and 17% (n=11) formally demonstrated the learning object. In terms of supports provided, 33% (n=21) provided a worksheet, while 31% of the teachers (n=20) created a set of guiding questions. Thirty-nine percent (n=25) of the teachers chose to discuss the learning object after it had been used.

Data sources

Learning Object Evaluation Metric (LOEM)
The original version of the LOEM had 29 items (see Appendix A at Kay and Knaack (2008c) for details) based on a thorough review of learning object evaluation criteria (see Table 1). Twelve items were excluded from the scale because of (a) insignificant correlations with student evaluations, student performance, and teacher evaluations and (b) insufficient fit into principal component analysis. The final version of the LOEM consisted of four constructs that were established from a detailed review of the literature: *Interactivity, Design, Engagement* and *Usability*. Note that the "content" construct supported by previous research (see Table 1) did not emerge as a significant factor. It is conceivable that teachers filtered "content" issues when they selected a learning object for their class. In other words, it is unlikely that they would select a learning object that did not have the correct content and scope. Consequently, content may have had a negligible influence. A detailed description of the scoring scheme for each item is presented in Appendix B (Kay & Knaack, 2008b)

Variables for assessing validity - students
Four dependent variables were chosen to assess validity of the LOEM from the perspective of the student: learning, quality, engagement, and performance. *Learning* referred to a student's self assessment of how much a learning object helped them learn. *Quality* was determined by student perceptions of the quality of the learning object. *Engagement* referred to student ratings of how engaging or motivating a learning object was. *Student performance* was determined by calculating the percent difference between pre- and post-test created by each teacher based on content of the learning object used in class.

Student self assessment of learning, quality and engagement was collected using the *Learning Object Evaluation Scale for Students* (LOES-S). These constructs were selected based on a detailed review of the learning object literature over the past 10 years (Kay & Knaack, 2007). The scale showed good reliability (0.78 to 0.89), face validity, construct validity, convergent validity and predictive validity.

Variables for *assessing validity - teachers*
Three dependent variables were chosen to assess validity of the LOEM from the perspective of the teacher: learning, quality and engagement. After using a learning object, each teacher completed the *Learning Object Evaluation Scale for Teachers* (LOES-T) to determine his/her perceptions of (a) how much their students learned (learning construct), (b) the quality of the learning object (quality construct), and (c) how much their students were engaged with the learning object (engagement construct). Data from the LOES-T showed low to moderate reliability (0.63 for learning construct, 0.69 for learning object quality construct, 0.84 for engagement construct), and good construct validity using a principal components factor analysis. See Kay & Knaack (2007b) for a detailed analysis of the teacher based learning object scale.

Data analysis

A series of analyses were run to assess the reliability and validity of the LOEM. These included:

1. internal reliability estimates (reliability);
2. a principal component factor analysis for learning object evaluation metric (LOEM) (construct validity);
3. correlations among LOEM constructs (construct validity);
4. correlation between LOEM and student evaluations (convergent validity);
5. correlation between LOEM and teacher evaluations (convergent validity); and
6. correlation between LOEM and student performance (predictive validity).

Results

Internal reliability

The Cronbach's internal reliability estimates for the four LOEM constructs were 0.70 (Interactivity), 0.74 (Design), 0.77 (Engagement), and 0.80 (Usability) – see Table 2. These moderate values are acceptable for measures in the social sciences (Kline, 1999; Nunally, 1978; see also Kay & Knaack, in press).

Table 2: Description of learning object evaluation metric

Scale	No. items	Possible range	Actual range observed	Mean (SD)	Internal reliability
Interactivity	3	3 to 9	3 to 9	6.0 (1.7)	$r = 0.70$
Design	4	4 to 12	4 to 12	9.3 (2.1)	$r = 0.74$
Engagement	5	5 to 15	5 to 15	9.4 (2.8)	$r = 0.77$
Usability	5	5 to 15	5 to 15	10.3 (2.7)	$r = 0.80$

Construct validity

Principal component analysis
A principal component analysis was done to explore whether the four LOEM constructs (interactivity, design, engagement, and usability) were distinct factors. Since all communalities were above 0.4 (Stevens, 1992), the principal component analysis was deemed an appropriate exploratory method (Guadagnoli & Velicer, 1988). An orthogonal varimax rotation was used because it simplifies the interpretation of the data (Field, 2005). The Kaiser-Meyer-Olkin measure of sampling adequacy (0.832) and Bartlett's test of sphericity (p <.001) indicated that the sample size was acceptable.

Table 3: Varimax rotated factor loadings on learning object evaluation metric (LOEM)

Scale item		Factor 1	Factor 2	Factor 3	Factor 4
Interactivity	Meaningful interactions	.799			
	Overall control	.723			
	Multimedia adds learning value	.645		.531	
Design	Consistency		.802		
	Layout		.702		
	Labelling		.633		
	Readability (look of text)		.574		
Engagement	Quality of feedback			.691	
	Attractive		.552	.682	
	Graphics			.659	
	Amount of multimedia			.656	
	Motivating			.581	
Usability	Natural to use		.538		.501
	Orientation				.794
	Navigational cues				.785
	Instructions				.559
	Appropriate language level			.519	.538
Factor	Eigenvalue	% of Variance	Cumulative %		
1	5.77	34.0	34.0		
2	2.12	12.5	46.5		
3	1.73	10.2	56.7		
4	1.21	7.1	63.8		

The principal components analysis extracted four factors (Table 3). The resulting rotation corresponded well with the proposed LOEM constructs with several exceptions. Factor 2, the design construct, included the four predicted scale items, but also showed relatively high loadings on 'attractiveness' (design construct) and 'natural to use' (usability construct). Factor 3, engagement, showed the highest loadings on the five predicted scale items, although 'multimedia adding learning value' (interactivity construct) and 'appropriate language level' (usability construct) items scored high as well. Overall, the resulting structure fit the proposed design model well.

Correlations among LOEM constructs

Correlations among the four LOEM constructs (interactivity, design, engagement, and usability) were significant, but small enough to support the assumption that each construct measured was distinct (Table 4).

Table 4: Correlations among learning object evaluation metric (LOEM) constructs

	Interactivity	Design	Engagement	Usability
Interactivity	1.00	0.31 *	0.46 *	0.42 *
Design		1.00	0.53 *	0.61 *
Engagement			1.00	0.39 *
Usability				1.00

* $p < .001$ (2-tailed)

Convergent validity

Correlation between LOEM and student evaluation (LOES-S)

Correlations among the four LOEM constructs and three LOES-S constructs were calculated to determine convergent validity. The LOEM interactivity construct was not significantly correlated with student perceptions of learning, quality, or engagement. However, the remaining three LOEM constructs (design, engagement, and usability) were significantly correlated with all three student evaluation constructs (Table 5).

Table 5: Correlations among LOEM and student evaluation constructs (LOES-S)

	S-Learn	S-Quality	S-Engagement
Interactivity	0.06	0.08	-0.08
Design	0.25 ($p < .001$, 2-tailed)	0.30 ($p < .001$, 2-tailed)	0.20 ($p < .01$, 2-tailed)
Engagement	0.24 ($p < .005$, 2-tailed)	0.32 ($p < .001$, 2-tailed)	0.27 ($p < .001$, 2-tailed)
Usability	0.28 ($p < .001$, 2-tailed)	0.27 ($p < .001$, 2-tailed)	0.15 ($p < .05$, 2-tailed)

Correlation between LOEM and teacher evaluation (LOES-T)
Correlations among the four LOEM constructs and three LOES-T constructs were calculated to determine convergent validity. Both the interactivity and usability constructs were significantly correlated with teachers' evaluation of learning object quality, but not learning or engagement. The design construct was significantly correlated with teachers' assessment of learning and quality, but not engagement. Finally, the engagement construct was significantly correlated with teachers' perceptions of learning, quality, and engagement with respect to learning objects (Table 6).

Table 6: Correlations among LOEM and teacher evaluation constructs (LOES-T)

	T-Learn	T-Quality	T-Engagement
Interactivity	0.06	0.19 ($p < .005$, 2-tailed)	-0.10
Design	0.15 ($p < .05$, 2-tailed)	0.19 ($p < .01$, 2-tailed)	0.06
Engagement	0.33 ($p < .001$, 2-tailed)	0.20 ($p < .05$, 2-tailed)	0.29 ($p < .001$, 2-tailed)
Usability	0.12	0.40 ($p < .001$, 2-tailed)	-0.13

Predictive validity

Correlation between LOEM constructs and student performance
One would predict that high scores in the LOEM constructs should result in increased gains in learning performance associated with the specific learning object in question. Mean class learning performance (percent change from the pre- to the post-tests) was significantly and positively correlated with the interactivity (r =.28; $p < .005$, n=119), design (r =.43; $p < .001$, n=116), engagement (r =.37; $p < .005$, n=84), and usability constructs (r =.40; $p < .001$, n=120).

Discussion

The purpose of the following study was to develop and assess a multi-component model for evaluating learning objects, based on an extensive review of criteria proposed by learning object theorists (see Table 1). Key issues that were targeted included sample population, range of learning objects assessed, theoretically grounded evaluation criteria, reliability, validity, using formal statistics where applicable, and student performance. Each of these issues will be discussed.

Sample population and range of learning objects

The population in this study was a large, diverse, sample of middle and secondary school students spread out over three school districts and 15 schools. This type of sampling is needed to build on previous small scale research efforts and establish generalisations.

The large number of learning objects tested is a significant departure from previous studies and offers evidence to suggest that the usefulness of the LOEM extends beyond a single learning object. While it is beyond the scope of this paper to compare specific types of learning objects used, it is reasonable to assume that the LOEM is a credible evaluation tool for a wide range of these learning tools. However, it should be noted that most of the learning objects were either mathematics or science based. Different results might have been obtained for other subject areas.

Reliability

The internal reliability estimates for the learning object constructs in the LOEM were acceptable (Kline, 1999; Nunally, 1978), but not exceptional. However, inter-rater reliability for assessing constructs was high, ranging from 94 to 96 percent. More work may be needed to increase reliability for interactivity and design constructs. One suggestion might be to revisit items using pilot interviews to determine sources of variability. Only two of the 24 formal evaluation studies reviewed for this paper offered reliability statistics (Baser, 2006; Kay & Knaack, 2005), yet it is argued that reliability is a fundamental element of any evaluation tool and should be calculated for future research studies, if the sample size permits.

Validity

Two of the 24 studies reviewed for this paper offered validity estimates (Kay & Knaack, 2007; Nurmi & Jaakkola, 2006a). However, it is critical to address validity in any learning object evaluation tool. Three types of validity were considered for this paper: construct, convergent, and predictive.

Construct validity

The principal components analysis revealed four relatively distinct learning object constructs (interactivity, design, engagement, and usability) that were consistent with the criteria (see Appendix A at Kay & Knaack, 2008a) proposed by previous learning object theorists. It should be noted that the majority of items that did not fit into the factor analysis (see Appendix A at Kay & Knaack, 2008a) tended to focus on basic functioning of a learning object (e.g. loading time, quality of video and sound), quality of content, or instructional supports (e.g. help instructions or information about a learning object). Since the teachers in this study selected their own learning objects, it is speculated that they may have filtered out basic problems in a learning object before selecting it. It is reasonable to assume that teachers would choose learning objects that loaded relatively quickly and had good multimedia quality, accurate content, and effective instructional supports. In essence, they were filtering out these variables and that is why they may not have loaded on the four construct model that emerged.

While the factors were relatively distinct, some items loaded on more than one factor construct. For example, learning objects that had ‘multimedia that added learning value’ loaded on both interactivity and engagement. These exceptions may indicate that defining discrete components is a complex issue and that while the overall structure is consistent with previous research, conceptual overlap may exist among interactivity, design, engagement, and usability. This conclusion is partially supported by the fact that correlations among learning object constructs were significant, but not too high, indicating that interactivity, design, engagement, and usability constructs were related but distinct factors.

Convergent validity
Convergent validity was supported by two tests. First, student estimates of learning, quality and engagement were significantly correlated with three of the four constructs in this model (design, engagement, and usability). According to the current model, well designed, engaging, easy to use learning objects correlate moderately well with student perceptions of the learning objects.

The second test of convergent validity was to match teacher perceptions of learning, quality, and engagement with the four concept model observed in this study. Overall, teacher ratings of learning, quality, and engagement correlated significantly with the interactivity, design, engagement, and usability constructs, with few exceptions.

Predictive validity
It is reasonable to predict that learning objects that are rated highly in terms of interactivity, design, engagement and/or usability should result in better learning performance. All four constructs showed positive and significant correlations with student performance.

Summary

The LOEM scale for evaluating learning objects was developed from a detailed list of criteria gleaned from a methodical review of the literature (Table 1). A large, diverse sample was used to test the LOEM multi-component model. A principal components analysis revealed four distinct constructs: interactivity, design, engagement, and usability. These four constructs demonstrated good internal and inter-rater reliability. They also correlated significantly with student and teacher perceptions of learning, quality, and engagement. Finally, all four constructs were significantly correlated with student learning performance. It is reasonable to conclude that the LOEM is reliable, valid, and effective approach to evaluating the effectiveness of learning objects in middle and secondary schools.

Implications for education

The main purpose for this paper was to develop a reliable, valid model tool for assessing learning objects. Ultimately, though, the goal of the LOEM is to help educators increase the pedagogical impact of technology in their classrooms. The LOEM can offer helpful guidance to educators who wish to select effective learning objects. The items presented in Appendix B (see Kay & Knaack, 2008b) offer practical guidance on key features to focus on when selecting a learning object. Learning objects that have good interactivity, clear design, high engagement, easy to use features, and high quality multimedia are most likely to lead to higher teacher and student satisfaction, as well as increased student performance.

One of the main challenges for educators is finding a suitable learning object in a reasonable amount of time. The development of the LOEM scale is a first step to creating learning object evaluation system for identifying effective learning objects. After further testing, the LOEM could be used to rate a series of learning objects in order to reduce educator search time. A database of good learning objects would also help promote reusability. If educators are confident in the LOEM evaluations given, they would be more likely to reuse "tried and true" learning objects, rather than search for or create new ones.

Caveats and future research

This study was designed with careful attention paid to the details of theory and methodology. An attempt was made to develop a learning object evaluation tool that was sensitive to key issues researched over the past ten years. A multi-component, theoretically driven model was developed and tested on a large, diverse sample, using a wide range of learning objects. Nonetheless, there are several caveats that should be addressed to guide future research.

First, strategies chosen to incorporate learning objects in the classroom probably have an impact on effectiveness. For example, a learning object used exclusively as a motivational or demonstration tool, might not have as much impact as a learning object used to teach a new concept. An analysis of instructional strategies could offer additional understanding of how to use learning objects more effectively. Second, tests used to assess performance in this study, were created on an ad hoc basis by individual teachers. No effort was made to standardise measures or to assess reliability and validity. Higher quality learning performance tools should increase the precision of results collected. Third, the learning objects tested in this study focussed primarily on mathematics and science. Markedly different features may be important for other subject areas. Future research should look at more diverse learning objects with respect to subject area. Fourth, the model was used to examine learning objects for grades 4 to 12. More research is needed to see if this model works for higher education, where the class size, maturity and independence of the students, and learning goals may be decidedly different. Fifth, most of the teachers in this study were very comfortable with computers. The scale needs to be tested on a more diverse group of teachers with respect to computer self efficacy and attitudes.

Finally, while there is a good evidence to suggest that interactivity, design, engagement, and usability are key features in selecting learning objects, there is no indication from the results as to how these constructs interact with learning. A cognitive task analysis is needed to determine how each of these constructs contributes to the learning process. For example, individual students could be asked to think aloud while using a learning objects to gather a more profound understanding of how interactivity, design, engagement, and usability influence decision making and performance.

References

Adams, A., Lubega, J., Walmsley, S. & Williams, S. (2004). The effectiveness of assessment learning objects produced using pair programming. *Electronic Journal of e-Learning,* 2(2). http://www.ejel.org/volume-2/vol2-issue2/v2-i2-art1-adams.pdf

Agostinho, S., Bennett, S., Lockyer, L. & Harper, B. (2004). Developing a learning object metadata application profile based on LOM suitable for the Australian higher education market. *Australasian Journal of Educational Technology,* 20(2), 191-208. http://www.ascilite.org.au/ajet/ajet20/agostinho.html

Akpinar, Y. & Bal, V. (2006). Student tools supported by collaboratively authored tasks: The case of work learning unit. *Journal of Interactive Learning Research,* 17(2), 101-119.

Anderson, T.A. (2003). I object! Moving beyond learning objects to learning components. *Educational Technology*, 43(4), 19-24.

Baruque, L. B. & Melo, R. N. (2004). Learning theory and instructional design using learning objects. *Journal of Educational Multimedia and Hypermedia,* 13(4), 343-370.

Baser, M. (2006). Promoting conceptual change through active learning using open source software for physics simulations. *Australasian Journal of Educational Technology*, 22(3), 336-354. http://www.ascilite.org.au/ajet/ajet22/baser.html

Bennett, K. & McGee, P. (2005). Transformative power of the learning object debate. *Open Learning*, 20(1), 15-30.

Bradley, C. & Boyle, T. (2004). The design, development, and use of multimedia learning objects. *Journal of Educational Multimedia and Hypermedia*, 13(4), 371-389.

Brown, A. R. & Voltz, B. D. (2005). Elements of effective e-learning design. *The International Review of Research in Open and Distance Learning*, 6(1). http://www.irrodl.org/index.php/irrodl/article/view/217/300

Butson, R. (2003). Learning objects: Weapons of mass instruction. *British Journal of Educational Technology*, 34(5), 667-669.

Buzzetto-More, N.A. & Pinhey, K. (2006). Guidelines and standards for the development of fully online learning objects. *Interdisciplinary Journal of Knowledge and Learning Objects,* 2006(2), 96-104. http://ijklo.org/Volume2/v2p095-104Buzzetto.pdf

Caws, C., Friesen, N. & Beaudoin, M. (2006). A new learning object repository for language learning: Methods and possible outcomes. *Interdisciplinary Journal of Knowledge and Learning Objects,* 2006(2), 112-124. http://ijklo.org/Volume2/v2p111-124Caws.pdf

Clarke, O. & Bowe, L. (2006a). The Le@rning Federation and the Victorian Department of Education and Training trial of online curriculum content with Indigenous students. 1-14. [viewed 1 Oct 2008] http://www.thelearningfederation.edu.au/verve/_resources/tlf_detvic_indig_trial_mar06.pdf

Clarke, O., & Bowe, L. (2006b). The Le@rning Federation and the Victorian Department of Education and Training trial of online curriculum content with ESL students. 1-16. [viewed 1 Oct 2008] http://www.thelearningfederation.edu.au/verve/_resources/report_esl_final.pdf

Cochrane, T. (2005). Interactive *QuickTime*: Developing and evaluating multimedia learning objects to enhance both face-to-face and distance e-learning environments. *Interdisciplinary Journal of Knowledge and Learning Objects,* 1. http://ijklo.org/Volume1/v1p033-054Cochrane.pdf

Concannon, F., Flynn, A. & Campbell, M. (2005). What campus-based students think about the quality and benefits of e-learning. *British Journal of Educational Technology,* 36(3), 501-512.

Convertini, V.C., Albanese, D., Marengo, A., Marengo, V. & Scalera, M. (2006). The OSEL taxonomy for the classification of learning objects. *Interdisciplinary Journal of Knowledge and Learning Objects,* 2, 125-138. http://ijklo.org/Volume2/v2p125-138Convertini.pdf

Del Moral, E. & Cernea, D.A. (2005). Design and evaluate learning objects in the new framework of the semantic web. In A. Mendez-Vila, B. Gonzalez-Pereira, J. Mesa Gonzalez & J.A. Mesa Gonsalez (Eds), *Recent research developments in learning technologies* (1-5). Spain: Formatux. [verified 28 Oct 2008] http://www.formatex.org/micte2005/357.pdf

Deaudelin, C., Dussault, M. & Brodeur, M. (2003). Human-computer interaction: A review of the research on its affective and social aspects. *Canadian Journal of Learning and Technology*, 29(1), 89-110. http://www.cjlt.ca/index.php/cjlt/article/view/34/31

Downes, S. (2003). Design and reusability of learning objects in an academic context: A new economy of education? *USDLA Journal*, 17(1). [viewed 1 June 2007, verified 28 Oct 2008] http://www.usdla.org/html/journal/JAN03_Issue/article01.html

Field, A. (2005). *Discovering statistics using SPSS* (2nd edition) Thousand Oaks, CA: SAGE Publications.

Fourner-Viger, P. (2006). A cognitive and logic based model for building glass-box learning objects. *Interdisciplinary Journal of Knowledge and Learning Objects*, 2, 77-94. http://ijklo.org/Volume2/v2p077-094Fournier-Viger.pdf

Friesen, N. (2001). What are educational objects? *Interactive Learning Environments*, 9(3). 219-230.

Friesen, N. & Anderson, T. (2004). Interaction for lifelong learning. *British Journal of Educational Technology*, 35(6), 679-687.

Gadanidis, G., Sedig, K. & Liang, H. (2004). Designing online mathematical investigation. *Journal of Computers in Mathematics and Science Teaching*, 23(3), 275-298.

Gibbons, A. S., Nelson, J. & Richards, R. (2000). The nature and origin of instructional objects. In D. A. Wiley (Ed.), *The instructional use of learning objects: Online version*. [viewed 1 July 2005, verified 28 Oct 2008] http://reusability.org/read/chapters/gibbons.doc

Guadagnoli, E. & Velicer, W. (1988). On methods in the analysis of profile data. *Psychometrika*, 24, 95-112.

Haughey, M. & Muirhead, B. (2005). Evaluating learning objects for schools. *E-Journal of Instructional Science and Technology*, 8(1). http://www.ascilite.org.au/ajet/e-jist/docs/vol8_no1/fullpapers/eval_learnobjects_school.htm

Howard-Rose, D. & Harrigan, K. (2003). CLOE learning impact studies lite: Evaluating learning objects in nine Ontario university courses. [viewed 3 July 2007, not found 28 Oct 2008] http://cloe.on.ca/documents/merlotconference10.doc

Jonassen, D. H. (2006). On the role of concepts in learning and instructional design. *Educational Technology Research & Development*, 54(2), 177-196.

Kay, R. H. & Knaack, L. (2005). Developing learning objects for secondary school students: A multi-component model. *Interdisciplinary Journal of Knowledge and Learning Objects*, 1, 229-254.

Kay, R. H. & Knaack, L. (in press). Assessing learning, quality and engagement in learning objects: The learning object evaluation scale for students (LOES-S). *Educational Technology Research & Development*.

Kay, R. H. & Knaack. L. (2008a). Appendix A: Possible variables for learning object evaluation measures. [viewed 19 Oct 2008] http://faculty.uoit.ca/kay/papers/AppendixA.html

Kay, R. H. & Knaack. L. (2008b). Appendix B: Learning object evaluation metric. [viewed 19 Oct 2008] http://faculty.uoit.ca/kay/papers/AppendixB.html

Kay, R. H. & Knaack. L. (2008c). Appendix C: List of learning objects used in the study. [viewed 19 Oct 2008] http://faculty.uoit.ca/kay/papers/AppendixC.html

Kay, R. H. & Knaack, L. (2007a). Evaluating the learning in learning objects. *Open Learning*, 22(1), 5-28.

Kay, R. H. & Knaack, L. (2007b). Teacher evaluation of learning objects in middle and secondary school classrooms. Manuscript submitted for publication.

Kenny, R. F., Andrews, B. W., Vignola, M. V., Schilz, M. A. & Covert, J. (1999). Towards guidelines for the design of interactive multimedia instruction: Fostering the reflective decision-making of preservice teachers. *Journal of Technology and Teacher Education*, 7(1), 13-31.

Kline, P. (1999). *The handbook of psychological testing* (2nd edition). London: Routledge.

Koohang, A. & Du Plessis, J. (2004). Architecting usability properties in the e-learning instructional design process. *International Journal on E-Learning*, 3(3), 38-44.

Koppi, T., Bogle, L. & Lavitt, N. (2004). Institutional use of learning objects: Lessons learned and future directions. *Journal of Educational Multimedia and Hypermedia*, 13(4), 449-463.

Koppi, T., Bogle, L. & Bogle, M. (2005). Learning objects, repositories, sharing and reusability. *Open Learning*, 20(1), 83-91.

Krauss, F. & Ally, M. (2005). A study of the design and evaluation of a learning object and implications for content development. *Interdisciplinary Journal of Knowledge and Learning Objects*, 1. http://ijklo.org/Volume1/v1p001-022Krauss.pdf

Lim, C. P., Lee, S. L. & Richards, C. (2006). Developing interactive learning objects for a computing mathematics models. *International Journal on E-Learning*, 5(2), 221-244.

Lin, A. & Gregor, S. (2006). Designing websites for learning and enjoyment: A study of museum experiences. *International Review of Research in Open and Distance Learning*, 7(3), 1-21. http://www.irrodl.org/index.php/irrodl/article/view/364/739

Lopez-Morteo, G. & Lopez, G. (2007). Computer support for learning mathematics: A learning environment based on recreational learning objects. *Computers & Education*, 48(4), 618-641.

LORDEC (2008a). Learning Object Research Development and Evaluation Collaboratory - Collections. [viewed 19 Oct 2008] http://www.education.uoit.ca/lordec/collections.html

LORDEC (2008b). Learning Object Research Development and Evaluation Collaboratory - Use. [viewed 19 Oct 2008] http://www.education.uoit.ca/lordec/lo_use.html

MacDonald, C. J., Stodel, E., Thompson, T. L., Muirhead, B., Hinton, C., Carson, B. & Banit, E. (2005). Addressing the eLearning contradiction: A collaborative approach for developing a conceptual framework learning object. *Interdisciplinary Journal of Knowledge and Learning Objects*, 1. http://ijklo.org/Volume1/v1p079-098McDonald.pdf

Malcolm, M. (2005). The exercise of the object: Issues in resource reusability and reuse. *British Journal of Educational Technology*, 26(1), 33-41.

Maslowski, R. & Visscher, A. J. (1999). Formative evaluation in educational computing research and development. *Journal of Research on Computing in Education*, 32(2), 239-255.

McCormick, R. & Li, N. (2005). An evaluation of European learning objects in use. *Learning, Media and Technology*, 31(3), 213-231.

McGreal, R. (2004). Learning objects: A practical definition. *International Journal of Instructional Technology and Distance Learning*, 1(9). http://www.itdl.org/Journal/Sep_04/article02.htm

McGreal, R., Anderson, T., Babin, G., Downes, S., Friesen, N., Harrigan, K., et al. (2004). EduSource: Canada's learning object repository network. *International Journal of Instructional Technology and Distance Learning*, 1(3). http://www.itdl.org/Journal/Mar_04/article01.htm

Metros, S. E. (2005). Visualizing knowledge in new educational environments: a course on learning objects. *Open Learning*, 20(1), 93-102.

Muzio, J. A., Heins, T. & Mundell, R. (2002). Experiences with reusable e-learning objects from theory to practice. *The Internet and Higher Education,* 5(1), 21-34.

Nesbit, J. & Belfer, K. (2004). Collaborative evaluation of learning objects. In R. McGreal (Ed.), *Online education using learning objects*. (pp. 138-153). New York: RoutledgeFalmer.

Nesbit, J., Belfer, K. & Vargo, J. (2002). A convergent participation model for evaluation of learning objects. *Canadian Journal of Learning and Technology,* 28(3). http://www.cjlt.ca/index.php/cjlt/article/view/110/103

Nielson, J. (2003). Ten usability heuristics. [viewed 1 June 2007, verified 28 Oct 2008] http://www.useit.com/papers/heuristic/heuristic_list.html

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

Nurmi, S. & Jaakkola, T. (2005). Problems underlying the learning object approach. *International Journal of Instructional Technology and Distance Learning*, 2(11). http://www.itdl.org/Journal/Nov_05/article07.htm

Nurmi, S. & Jaakkola, T. (2006a). Effectiveness of learning objects in various instructional settings. *Learning, Media and Technology*, 31(3), 233-247.

Nurmi, S. & Jaakkola, T. (2006b). Promises and pitfall of learning objects. *Learning, Media, and Technology*, 31(3), 269-285.

Ohl, T. M. (2001). An interaction-centric learning model. *Journal of Educational Multimedia and Hypermedia,* 10(4), 311-332.

Oliver, R. & McLoughlin, C. (1999). Curriculum and learning-resources issues arising from the use of web-based course support systems. *International Journal of Educational Telecommunications*, 5(4), 419-435.

Parrish, P. E. (2004). The trouble with learning objects. *Educational Technology Research & Development,* 52(1), 49-67.

Reimer, K. & Moyer, P.S. (2005). Third-graders learning about fractions using virtual manipulatives: A classroom study. *Journal of Computers in Mathematics and Science Teaching,* 24(1*),* 5-25.

Schell, G. P., & Burns, M. (2002). A repository of e-learning objects for higher education. *e-Service Journal,* 1(2), 53-64.

Schoner, V., Buzza, D., Harrigan, K. & Strampel, K. (2005). Learning objects in use: 'Lite' assessment for field studies. *Journal of Online Learning and Teaching,* 1(1), 1-18. [verified 28 Oct 2008] http://jolt.merlot.org/documents/vol1_no1_schoner_001.pdf

Siqueira, S. W. M., Melo, R. N. & Braz, M. H. L. B. (2004). Increasing the semantics of learning objects. *International Journal of Computer Processing of Oriental Languages,* 17(1), 27-39.

Sosteric, M. & Hesemeier, S. (2002). When is a learning object not an object: A first step towards a theory of learning objects. *International Review of Research in Open and Distance Learning,* 3(2), 1-16. http://www.irrodl.org/index.php/irrodl/article/view/106/185

Sosteric, M & Hesemeier, S. (2004). A first step towards a theory of learning objects. In R. McGreal (Ed.), *Online education using learning objects* (pp. 43-58). London: Routledge Falmer.

Stevens, J. P. (1992). *Applied multivariate statistics for the social science applications* (2nd edition). Hillsdale, NJ: Erlbaum.

Wiley, D., Waters, S., Dawson, D., Lambert, B., Barclay, M. & Wade, D. (2004). Overcoming the limitations of learning objects. *Journal of Educational Multimedia and Hypermedia*, 13(4), 507-521.

Williams, D. D. (2000). Evaluation of learning objects and instruction using learning objects. In D. A. Wiley (Ed.), The instructional use of learning objects: Online version. [viewed 1 July 2005] http://reusability.org/read/chapters/williams.doc

Van Merrienboer, J. J. G. & Ayres, P. (2005). Research on cognitive load theory and its design implications for e-learning. *Education Theory, Research and Development*, 53(3), 1042-1629.

Van Zele, E., Vandaele, P., Botteldooren, D. & Lenaerts, J. (2003). Implementation and evaluation of a course concept based on reusable learning objects. *Journal of Educational Computing and Research*, 28(4), 355-372.

Vargo, J., Nesbit, J. C., Belfer, K. & Archambault, A. (2002). Learning object evaluation: Computer mediated collaboration and inter-rater reliability. *International Journal of Computers and Applications*, 25(3), 1-8.

Dr Robin H. Kay
University of Ontario Institute of Technology
Faculty of Education
2000 Simcoe St. North, Oshawa, Ontario L1H 7L7, Canada
Email: Robin.Kay@uoit.ca Web: http://faculty.uoit.ca/kay/home/

Dr Liesel Knaack
University of Ontario Institute of Technology
Faculty of Education
Email: liesel.knaack@uoit.ca