Assessing Learning, Quality and Engagement in Learning Objects:

The Learning Object Evaluation Scale for Students (LOES-S)

Paper Submitted for Publication at ETRD (Nov 2007)
Abstract

Research on the impact, effectiveness, and usefulness of learning objects is limited, partially because comprehensive, theoretically-based, reliable, and valid evaluation tools are scarce, particularly in the K-12 environment. The purpose of the following study was to investigate a learning object evaluation scale for students (LOES-S) based on three key constructs emphasized in 10 years of learning object research: learning, quality or instructional design, and engagement. Tested on over 1100 middle and secondary school students, the LOES-S showed good internal reliability, face validity, construct validity, convergent validity and predictive validity.

Keywords: evaluate; assess; quality; scale; secondary school; middle school; learning object
Assessing Learning, Quality and Engagement in Learning Objects:

The Learning Object Evaluation Scale for Students (LOES-S)

Overview

The design, development, reuse, accessibility, and use of learning objects has been examined in some detail for almost 10 years (Kay & Knaack, 2007), however, research on the impact, effectiveness, and usefulness of learning objects is limited (Kay & Knaack, 2005; Nurmi & Jaakkola, 2005, 2006a, 2006b, Sosteric & Hesemeirer, 2004). While the challenge of developing an effective, reliable, and valid evaluation system is formidable (Kay & Knaack, 2005; Nesbit & Belfer, 2004), assessing effectiveness is critical if learning objects are to be considered a viable educational tool.

To date, the evaluation of learning objects has taken place predominantly at the development and design phase (Adams, Lubega, Walmsley & Williams, 2004; Bradeley & Boyle, 2004; Maslowski & Visscher, 1999; Vargo, Nesbit, Belfer & Archambault, 2002; Williams, 2000). This kind of formative analysis, done while learning objects are created, is useful for developing easy to use learning objects, but the voice of the end user, the student who uses the learning object, is relatively silent.

A limited number of repositories have content experts, often educators, evaluate the quality of learning objects (Cafolla, 2006; Krauss & Ally, 2005; Schell & Burns, 2002) after they have been develop. However, the number of evaluators is usually limited, the assessors have limited background in instructional design, and the end user does not enter the feedback loop in a significant way.

Until recently, learning objects were solely used in higher education, therefore the majority of learning object evaluation has taken place in this domain (Haughey & Muirhead,
Increased use of learning objects in the K-12 domain (e.g., Brush & Saye, 2001; Clarke & Bowe, 2006a, 2006b; Kay & Knaack, 2005; Lopez-Morteo & Lopez, 2007; Liu & Bera, 2005; Nurmi & Jaakkola, 2006a) demands that the focus of evaluation shift, at least in part, to the needs of middle and secondary school students.

The purpose of the current study is to examine a student–based learning object evaluation tool designed to look at learning, quality, and engagement of the end user while using learning objects in middle and secondary school classrooms.

Literature Review

Theory Underlying the Evaluation of Learning Objects

While current methods used to evaluate learning objects are somewhat limited with respect to underlying theory (e.g., Buzetto-More & Pinhey, 2006; Gadanidis, Sedig, & Liang, 2004; Koohang & Plessis, 2004; McGreal et al., 2004; Schoner et al., 2005), considerable speculation and discussion has taken place on the necessary attributes for developing effective assessment tools. Three key themes have emerged from these discussions: learning, quality or instructional design, and engagement.

Learning. The learning objects field has addressed technical, instructional design issues in evaluation far more than those based on pedagogy (Alonso, Lopez, Manrique, & Vines, 2005; Jonassen, 2006; Kay & Knaack, 2005). This approach has resulted in a model of learning that is dated and largely behaviouristic – content is presented, students are asked questions, evaluated and rewarded based on the content they remember (Friesen, 2004; Krauss & Ally, 2005; Nurmi & Jaakkola, 2006b). For the past 20 years, though, research in cognitive science suggests that students need to construct knowledge and actively participate in the learning process (e.g., Albanese & Mitchell, 1993; Bruner, 1983, 1986; Brown & Palinscar, 1989; Chi & Bassock,
and within the last five years, several learning object theorists have advocated the use of a more constructivist-based metric (Baser, 2005; Convertini, Albanese, Marengo, Marengo, & Scalera, M., 2005; Gadanidis et al., 2004).

One way of assessing constructivism is by measuring the amount and quality of interactivity in a learning object. While there is a tendency to view all interactivity as wonderful, considerable debate reigns on narrowing down the key elements of good interaction (Ohl, 2001). Van Marnienboer & Ayres (2005) speculate that interaction requiring a student to manipulate digital learning materials may be more motivating and stimulating. Lim, Lee, & Richards (2006) have proposed six different levels of interactivity including mouse pointing and clicking, linear navigation, hierarchical navigation, interacting with help, program generated questions, and constructing or manipulating. Finally, Oliver & McLoughlin (1999) argue that, ideally, students who use learning objects should be making reasoned actions, engaging in personal meaning making, and integrating knowledge.

To date, constructivism and interactivity have not been systematically integrated into the learning object evaluation process. However, there is some evidence that students believe learning features of a learning object are more important than the technical features (Kay & Knaack, 2005; 2007).

**Quality (Instructional Design).** For the purpose of this paper, the quality of a learning object refers to technical, design issues focusing on usability, as opposed to the learning issues discussed above. Evaluating the quality of learning objects is based on a wealth of research looking at the instructional design of digital materials and includes the following features: organization & layout (Calvi, 1997; Madhumita, 1995), learner control (Druin et al. 1999; Hanna, Risden, Czerwinski, & Alexander, 1999; Kennedy & McNaught, 1997), multimedia
the form of animation, graphics, and audio (Gadanidis, Gadanidis, & Schindler, 2003; Sedig & Liang, 2006), clear instructions & help features (Acovelli et al., 1997; Jones, Farquhar, & Surry, 1995; Kennedy & McNaught, 1997), feedback and assessment (Kramarski & Zeichner, 2001; Zammit, 2000), and theme (Akpibar & Hartley, 1996; Harp & Mayer, 1998). In spite of this well researched list of qualities that have been reported to effect software usability, summative evaluation tools filled in by users or students rarely address instructional design qualities of learning objects. It is more typical to collect open-ended, informal feedback without reference to specific instructional design characteristics that might enhance or reduce learning performance (e.g., Bradley & Boyle, 2004; Kenny, Andrews, Vignola, Schilz, & Covert, 1999; Krauss & Ally, 2005).

Cognitive load theory (Chandler & Sweller, 1991; Kester, Lehnen, Van Gerven, & Kirschner, 2006; Sweller, 1988, Sweller, van Merriënhoer, & Paas, 1998) has been used to organize and explain the potential impact that the features of a learning object can have on performance. The main premise is that a typical user wants to minimize extraneous cognitive load (engaging in processes that are not beneficial to learning) and optimize germane cognitive load (engaging in processes that help to solve the problem at hand). Therefore, if the quality or instructional design of a learning object is sufficiently weak in one or more areas, the user spends more time on trying to use the object than on learning the concept at hand. Because the quality of learning objects is rarely addressed in the literature, little is known about how learning object design features affect cognitive load and ultimately how much is learned by the end user.

Engagement. A number of authors believe that a high level of engagement or motivation is necessary for a learning object to be successful. Lin & Gregor (2006) suggest that engagement, positive affect, and personal fulfilment are key factors in the evaluation process.
Oliver & McLoughlin (1999) add that self-efficacy is critical to promoting engagement in learning objects. Van Marrienboer & Ayres (2005) note that lower task involvement, as a result of reduced motivation, can result in a lower investment of cognitive effort. In summary, it is important to consider the degree to which a learning object engages students when evaluating effectiveness.

Previous Approaches to Evaluating Learning Objects

Considerable effort has been directed toward the evaluation of learning objects as they are being created (Adams et al. 2005; Bradley & Boyle, 2004; Cochrane, 2005; MacDonald et al., 2005; Nesbit, Belfer, & Vargo, 2002; Vargo et al., 2003). Also known as formative assessment, this approach to evaluation typically involves a small number of participants being asked to test and use a learning object throughout the development process. Cochrane (2005) provides a good example of how this kind of evaluation model works where feedback is solicited from small groups at regular intervals during the development process. While formative evaluation is necessary for the development of learning objects that are well designed from a usability standpoint, this type of assessment does not address how well the learning object works in a real-world educational environment with actual students.

Qualitative analysis of learning objects is also prevalent in the evaluation literature in the form of interviews (Bradley & Boyle, 2004; Kenny et al., 1999; Lin & Gregor, 2006), written comments (Kay & Knaack, 2005; Kenny et al., 1999; Krauss & Ally, 2005), email responses (Bradley & Boyle, 2004; MacDonald et al., 2005) and think-aloud protocols (Holzinger, 2004; Krauss & Ally, 2005). The majority of studies using a qualitative approach rely almost exclusively on descriptive data and anecdotal reports to assess the merits of learning objects. The reliability and validity of these informal qualitative observations are questionable.
Quantitative efforts to evaluate learning objects have incorporated surveys (Bradley & Boyle, 2004; Howard-Rose & Harrigan, 2003; Krauss & Ally, 2005), performance data (Adams et al., 2005; Bradley & Boyle, 2004; Nurmi & Jaakola, 2006a), and use statistics (Bradley & Boyle, 2004; Kenny et al., 1999). The main concerns with the quantitative measures used to date are a lack of theory underlying measures and the absence of reliability and validity estimates.

A common practice employed to evaluate learning objects is to use multiple assessment tools (Bradley & Boyle, 2004; Brown & Voltz, 2005; Cochrane, 2005; Kenny et al., 1999; Krauss & Ally, 2005; Nesbit & Belfer, 2004; Schell & Burns, 2002; Schoner et al., 2005; Van Zele, Vandaele, Botteldooren, & Lenaerts, 2003). This approach, which leads to triangulation of data analysis, should be encouraged, however, the multitude of constructs that have evolved to date do not provide a coherent model for understanding what factors contribute to the effectiveness of learning objects.

Methodological Issues

At least six key observations are noteworthy with respect to methods used to evaluate learning objects. First, a wide range of learning objects have been examined including drill-and-practice assessment tools (Adams et al., 2004) or tutorials (Nurmi & Jaakkola, 2006a), video case studies or supports (Kenny et al., 1999; MacDonald et al., 2005), general web-based multimedia resources (Van Zele et al., 2003), and self-contained interactive tools in a specific content area (Bradley & Boyle, 2004; Cochrane, 2005). The content and design of a learning object need to be considered when examining quality and learning outcomes. For example, Cochrane (2005) compared a series of four learning objects based on general impressions of reusability, interactivity, and pedagogy and found that different groups valued different areas.
As well, Nurmi & Jaakkola (2004) compared drill-and-practice versus interactive learning objects and found the latter to be significantly more effective in improving overall performance.

Second, even though a wide range of learning objects exist, the majority of evaluation papers focus on a single learning object (Adams et al., 2004; Bradley & Boyle, 2004; Kenny et al., 1999; Krauss & Ally, 2005; MacDonald et al., 2003). It is difficult to determine whether the evaluation tools used in one study generalize to the full range of learning objects that are available.

Third, while the number of studies focusing on the K to 12 population has increased recently (e.g., Brush & Saye, 2001; Clarke & Bowe, 2006a, 2006b; Kay & Knaack, 2005; Lopez-Morteo & Lopez, 2007; Liu & Bera, 2005; Nurmi & Jaakkola, 2006a), most evaluation of learning objects has been done in the domain of higher education.

Fourth, sample populations tested in many studies have been noticeably small and poorly described (e.g., Adams et al., 2004; Cochrane, 2005; Krauss & Ally, 2005; MacDonald et al., 2005; Van Zele et al., 2003) making it challenging to extend any conclusions to a larger population.

Fifth, while most evaluation studies reported that students benefited from using learning objects, the evidence is based on loosely designed assessment tools with no validity or reliability (Bradley & Boyle, 2004; Howard-Rose & Harrigan, 2003; Krauss & Ally, 2005; Kenny et al., 1999; Lopez-Morteo & Lopez, 2007; Schoner et al., 2005; Vacik et al., 2006; Van Zele et al., 2003; Vargo et al., 2003). As well, very few evaluation studies (e.g., Kenny et al., 1999; Kay & Knaack, 2007; Van Zele et al., 2003) use formal statistics. The lack of reliability and validity of evaluation tools combined with an absence of statistical rigour reduce confidence in the results presented to date.
Finally, a promising trend in learning object evaluation research is the inclusion of performance measures (e.g., Adams et al., 2004; Bradley & Boyle, 2004; Docherty et al., 2005; MacDonald et al., 2005; Nurmi & Jaakkola, 2006a). Until recently, there has been little evidence to support the usefulness or pedagogical impact of learning objects. The next step is to refine current evaluation tools to determine which specific qualities of learning objects influence performance.

In summary, previous methods used to evaluate learning objects have offered extensive descriptive and anecdotal evaluations of single learning objects, but are limited with respect to sample size, representative populations, reliability and validity of data collection tools, and the use of formal statistics. Recent evaluation efforts to incorporate learning performance should be encouraged in order to advance knowledge of learning object features that may influence learning.

Current Approach to Evaluating Learning Objects

Definition of Learning Objects. In order to develop a clear, effective metric, it is necessary to establish an operational definition of a “learning object”. Original definitions focussed on technological issues such accessibility, adaptability, the effective use of metadata, reusability, and standardization (e.g., Downes, 2003; Littlejohn, 2003; Koppi, Bogle, & Bogle, 2005; Muzio, Heins, & Mundell, 2002; Nurmi & Jaakola, 2006b; Parrish, 2004; Siqueira, Melo, & Braz, 2004). More recently, a number of researchers are emphasizing learning qualities such as quality of interaction and degree to which the learner actively constructs knowledge (Baruque & Melo, 2004; Bennett & McGee, 2005; Bradley & Boyle, 2004; Caws, Friesen, & Beaudoin, 2006; Chenail, 2004; Cochrane, 2005; McGreal, 2004; Kay & Knaack, 2007; 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, evaluations tools focusing on learning are noticeably absent (Kay and Knaack, 2007). In order to address a clear gap in the literature on evaluating learning objects, a pedagogically focused definition of learning objects has been adopted for the current study. Learning objects are defined as “interactive web-based tools that support the learning of specific concepts by enhancing, amplifying, and guiding the cognitive processes of learners”.

Theoretical Model. The model used to support the evaluation tools in this study was based on a (a) thorough review of the literature on learning objects (see above) and (b) recent feedback from a similar evaluation tool developed by Kay & Knaack (2007). Consequently, three key constructs were developed for the quantitative survey and included learning, quality, and engagement (see Appendix A). The learning construct referred to a student’s perception of how much he/she learned from using the learning object. The quality construct referred to the design of the learning object and included the following key instructional design features identified by Kay & Knaack (2007): help features, clarity of instructions, ease of use, and organization. Finally, the engagement construct examined how involved a student was with respect to using a learning object. Estimates of all three constructs were supplemented by written comments that students made about what they liked and did not like about the learning object. The qualitative coding rubric used in the current study incorporated learning benefits and a full range of instructional design features (see Table 2). Finally, learning performance was incorporated into the evaluation system.
Purpose

The purpose of this study was to explore a comprehensive learning-based approach for evaluating learning objects. Based on a detailed review of studies looking at the evaluation of learning objects, the following steps were followed:

1. a large, diverse, sample was used;
2. a wide range of learning objects were tested;
3. the design of the evaluation tools was based on a thorough review and categorization of the learning object literature and instructional design research;
4. reliability and validity estimates were calculated;
5. formal statistics were used where applicable;
6. both qualitative and quantitative data were collected, systematically coded, and analysed;
7. measure of learning performance were included; and
8. evaluation criteria focussed on the end user perceptions and not those of the learning object designers.

Method

Sample

Students. The student 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 separate boards of education, six middle schools, 15 secondary schools, and 33 different classrooms. The students were selected through convenience sampling and had to obtain signed parental permission to participate.
Teachers. The teacher sample consisted of 33 teachers (12 males, 21 females), with 0.5 to 33 years of teaching experience ($M = 9.0, SD = 8.2$), from both middle ($n=6$) and secondary schools ($n=27$). Most teachers taught math ($n=16$) 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 ($n=29$). In spite of the high ability and positive attitude, only six of the teachers used computers in their classrooms more than once a month.

Learning Objects. A majority of teachers selected learning objects from a repository located at the LORDEC website (http://www.education.uoit.ca/lordec/collections.html), although several reported that they also used Google. A total of 44 unique learning objects were selected covering concepts in biology, Canadian history, chemistry, general science, geography, mathematics, and physics.

Procedure

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 http://www.education.uoit.ca/lordec/lo_use.html 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, however, 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.
Data Sources

Student survey. After using a learning object, students completed the Learning Object Evaluation Scale for Students (LOES-S) in Appendix A to determine their perception of (a) how much they learned (learning construct), (b) the quality of the learning object (quality construct), and (c) how much they were engaged with the learning object (engagement construct). Descriptive statistics for the LOES-S are presented in table 1.

Student comments. Students were asked to comment on what they liked and disliked about the learning object (Appendix A – questions 13 and 14). These qualitative items were organized according to the three main constructs identified in the literature review (learning, quality, and engagement) and analysed using the coding scheme provided in Table 2. This coding scheme (Kay & Knaack, 2007) was used to categorize 1922 student comments. Each comment was then rated on a five-point Likert scale (-2 = very negative, -1 = negative, 0 = neutral, 1 = positive, 2 = very positive). Two raters assessed all comments made by students and achieved inter-rater reliability of 99% on the categories and 100% on the ratings.

Student performance. Students completed a pre-test and pos-test created by each teacher based on the content of the learning object used in class. The measure was used to determine student performance.
**Teacher survey.** After using a learning object, each teacher completed the Learning Object Evaluation Scale for Teachers (LOES-T) to determine their perception 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). The LOES-T showed fair to moderate reliability and good construct validity (see Kay & Knaack, submitted for publication).

**Data Analysis**

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

1) internal reliability estimates (reliability);
2) a principal component factor analysis for Student Learning Object Evaluation Scale (LOES-S) (construct validity);
3) correlations among learning object evaluation constructs within the LOES-S scales (construct validity);
4) correlation between LOES-S and LOES-T constructs (convergent validity);
5) correlation between LOES-S and computer comfort level (convergent validity);
6) correlations between coded student comments and LOES-S constructs (convergent validity);
7) correlation between learning performance and LOES-S constructs (predictive validity);
Results

**Internal Reliability**

The internal reliability estimates for the LOES-S constructs were 0.89 (Learning), 0.84 (Quality), and 0.78 (Engagement) – see Table 2. These moderate to high values are acceptable for measures in the social sciences (Kline, 1999; Nunally, 1978).

**Construct Validity**

*Principal component analysis.* A principal components analysis was done to explore whether the three learning object constructs (learning, quality, and engagement) in the LOES-S formed three distinct factors. Since all communalities were above 0.4 (Stevens, 1992), the principal component analysis was deemed an appropriate exploratory method (Guadagnoli & Velicer, 1988). Orthogonal (varimax) and oblique (direct oblimin) rotations were used, given that the correlation among potential strategy combinations was unknown. These rotational methods produced identical factor combinations, so the results from the varimax rotation (using Kaiser normalization) are presented because they simplify the interpretation of the data (Field, 2005). The Kaiser-Meyer-Olkin measure of sampling adequacy (0.937) and Bartlett’s test of sphericity ($p < .001$) indicated that the sample size was acceptable.

The principal components analysis was set to extract three factors (Table 3). The resulting rotation corresponded well with the proposed learning object evaluation constructs with several exceptions. Factor 1, the learning construct, included the five predicted scale items, but also showed relatively high loadings on one of the quality construct items and two of the engagement construct items. Factor 3, engagement, showed the highest loadings on the predicted scale items, however, two learning construct items scored high as well. Overall, the structure was
consistent with previous research (Kay, 2007) and the proposed grouping of scale items listed in Appendix A.

Correlations among LOES-S constructs. Correlations among the three LOES-S constructs (learning, quality, and engagement) were significant, but small enough to support the assumption that each construct measured was distinct (Table 4).

Convergent Validity

Correlation between LOES-S and LOES-T constructs. All three LOES-S constructs were significantly correlated with all three LOES-T constructs. In other words, there was a certain degree of consistency between student and teacher evaluations of learning objects using the LOES-S and LOES-T scales (Table 5).

Correlation between student computer comfort level and LOES-S constructs. Computer comfort level based on a 3 item scale (Kay & Knaack, 2007) was significantly correlated with
the learning ($r = .27; p < .001$), quality ($r = .26; p < .001$), and engagement constructs ($r = .32; p < .001$). The more comfortable that a student was with the computers, the more likely he/she would rate learning, quality, and engagement of a learning object higher.

**Correlation between student comments and LOES-S constructs.** The LOES-S learning construct showed significant correlations with comments made by students about learning ($r = 0.27, p < .001$), challenge level ($r = 0.27, p < .001$), and visual aids ($r = 0.11, p < .005$). The LOES-S quality construct showed small but significant correlations with students perceptions of a learning object being easy to use ($r = 0.09, p < .05$), quality of help given ($r = 0.09, p < .05$), and the quality/amount of text in a learning object ($r = 0.09, p < .05$). Finally, the LOES-S engagement construct showed significant correlations with comments made by students about engagement ($r = 0.21, p < .001$).

**Predictive Validity**

**Correlation between learning performance and LOES-S constructs.** Learning performance (percent change from the pre to the post tests) for classes where a learning object was not used for review ($n=273$), was significantly and positively correlated with the learning ($r = .18; p < .01, n=254$), quality ($r = .22; p < .005, n=242$), and engagement constructs ($r = .10; p < .005, n=273$). In other words, higher scores on student perceptions of learning, learning object quality, and engagement were associated with higher scores in learning performance, although this effect is relatively small.

**Discussion**

The purpose of this study was to systematically investigate a student focussed approach for evaluating learning objects, based on three prominent themes that appeared in previous research: learning, quality, and engagement. Key issues addressed were sample population,
range of learning objects assessed, reliability, validity, using formal statistics where applicable, incorporating both qualitative and quantitative feedback, and student performance. Each of these issues will be discussed in turn.

Sample Population & 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 in order to provide in depth analysis and confidence regarding specific learning object features that affected learning.

The sizable number of learning objects tested is a significant departure from previous studies and offers evidence to suggest that the usefulness LOES-S 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 LOES-S is a credible evaluation tool for a wide range of these learning tools.

Reliability

The internal reliability estimates for the learning object constructs in the LOES-S were good (Kline, 1999; Nunally, 1978), as was the inter-rater reliability of the categories and ratings used to assess student comments. Less than 25% of the 25 formal evaluation studies reviewed for this paper (Baser, 2005; Kay & Knaack, 2005; Kong & Kwok, 2005; Liu & Bera, 2005; Vargo et al., 2003; Windschitl & Andre, 1998) offered reliability statistics, yet it argued that reliability is a fundamental aspect element of any evaluation tool and should be calculated for future research studies, if the sample size permits.

Validity
Only two of the 25 studies reviewed for this paper, offer validity estimates (Kay & Knaack, 2007; Nurmi & Jaakkola, 2006a), therefore it is important to address validity in any learning object evaluation tool. Four types of validity were considered for this paper: face, construct, convergent, and predictive.

**Face Validity**

Face validity was supported by (a) the close alignment between the three proposed LOES-S constructs (learning, quality, and engagement) and those features identified as important in comprehensive review of the literature, and (b) the coding of student comments which also matched the three LOES-S subscales. Aligning scale constructs with a systematic analysis of previous theory is critical if face validity is to be established.

**Construct Validity**

The principal components analysis revealed three relatively distinct learning object constructs that were consistent with the theoretical framework proposed by previous learning object researchers and instructional design specialists. However, three exceptions were noted. First, the feedback and graphics items (items 3 and 3 from Appendix A), designed to measure learning, overlapped with the engagement construct. Second, quality of help (item 6 in Appendix A) showed high communality with both learning and quality constructs. Third, theme and motivation (items 10 and 11 in Appendix A) showed high loading on both learning 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 learning, quality, and motivation. This conclusion is partially supported by the fact that correlations among learning object constructs were significant, but not too high, indicating that learning, quality, and engagement constructs were related but distinct.
Convergent Validity

Convergent validity was supported by three tests. First, correlations between student estimates of learning, quality, and engagement were significantly correlated with teacher estimates of these same constructs. The correlations were not overly high, but that might be expected given that teachers and students may have different perceptions on what constitutes learning, quality, and engagement. Therefore, while student and teacher constructs do converge, the modest correlations underline one of the main premises of this paper, namely the need for obtaining student input.

The second test of convergent validity looked at correlations among the three LOES-S constructs and student computer comfort level. It was predicted that students who were more comfortable with computers would rate learning objects more highly in terms of learning, quality, and motivation. This prediction was supported by the low, but significant correlations observed.

The third test for convergent validity was looked at by comparing survey construct scores and student comments. While the LOES-S learning construct was significantly correlated with student comments made about learning, and the LOES-S engagement construct correlated significantly with engagement comments, LOES-S quality construct showed a much weaker, but significant association with student assessment of quality. This result may be partially explained by the wide range of instructional qualities involved in a learning object, compared to the evaluation of learning and engagement. Increased variability invites complexity.

Predictive Validity

It is reasonable to predict that learning objects that are rated highly in terms of learning, quality, and/or engagement should result in better learning performance. In other words, if a
student perceives a learning object as a high quality, effective learning tool that is engaging, we would expect him/her to perform better in a pre-post test situation. Significant, but small, correlations between the percent change in pre-post test scores and the LOES-S learning, quality, and engagement constructs supported these predictions. One might expect these correlations to be higher if the LOES_S is to be an effective assessment tool, however, learning is a complex process involving numerous variables including instructional wrap, student ability in a subject area, student attitude toward a subject, gender, and time using the learning object. The challenges and problems of education are always more complex than technology can solve, therefore, modest correlations between key learning object constructs and learning performance are probably to be expected.

Implications for Education

The main purpose for this paper was to develop a reliable, valid student-based evaluation tool for assessing learning objects, however, there are several implications for education. First, it is prudent to gather student input when using learning objects and other technologies in the classroom. While teacher and student assessment of learning benefits, quality, and engagement are consistent with each other, they only share 20% common variance. It is through student feedback that these tools and the instructional wrap that supports them can be improved.

Second, the evaluation tool in this study offers some guidance on key features to focus on when selecting a learning object. Learning features such as good interactivity, clear feedback, and graphics or animations that support learning are desirable, as are design qualities such as effective help, clear instructions, transparency of use and organization. It is probably more challenging for an educator to understand what engages a student, however overall theme can impact positively or negatively on learning.
Finally, it is important to remember the low but significant correlations among student evaluations of learning, quality, and engagement and learning performance. No technology will transform the learning process. Learning objects are simply tools used in a complex educational environment where decisions on how to use these tools may have considerably more import than the actual tools themselves.

*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 three-prong model was developed and tested on a large, diverse sample, using a wide range of learning objects. Nonetheless, there several caveats that should be addressed to guide future research.

First, student ability was not examined and may have an impact on the success of any learning tool, let alone a learning object. In other words, students who like a subject may be more open to new ways of learning concepts, whereas students who are struggling may find the use of a learning object distracting and challenging. Assessing student ability might provide further clarity in future research endeavours.

Second, instructional wrap or strategies employed to incorporate learning objects in the classroom probably have an impact on the 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. While some effort was made to assess how the learning object was used (e.g., for review), a more detailed analysis of instructional wrap could offer additional understanding of learning object usefulness.
Third, more directed and through qualitative analysis is needed to assess the design qualities of learning objects. The open-ended approach used in the current study generated a wide range of responses and these responses could now be used to gather more in depth data. For example, students could be asked what features specifically supported and detracted from their learning and how they would design the learning object to be more effective.

Fourth, tests used to assess performance in this study, were created on an ad hoc basis, by individual teachers. No effort was made to standardize measures or to assess reliability and validity. Higher quality learning performance tools should increase the precision of results collected.

Finally, it would be extremely beneficial to compare systematic external evaluations of learning objects by experts with student evaluations and performance. We could then begin to assess whether design efforts and certain types of learning objects have their intended impact.
References


Chenail, R.J. (2004). When Disney meets the research park: Metaphors and models for engineering an online learning community of tomorrow. *Internet and Higher Education, 7* (2), 107-121.


economy of education? USDLA, 17(1). Retrieved June 1, 2007 from


Developing learning objects for secondary school students: A multi-component model.

Interdisciplinary Journal of Knowledge and Learning Objects, 2005 (1), 229-254.


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.


Table 1

Description of Student Learning Object Evaluation Scales (LOES-S)

<table>
<thead>
<tr>
<th>Scale</th>
<th>No. Items</th>
<th>Possible Range</th>
<th>Actual Range</th>
<th>Internal Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOES-S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learn</td>
<td>5</td>
<td>5 to 25</td>
<td>12.6 to 20.6</td>
<td>$r = 0.89$</td>
</tr>
<tr>
<td>Quality</td>
<td>4</td>
<td>4 to 20</td>
<td>10.9 to 18.3</td>
<td>$r = 0.84$</td>
</tr>
<tr>
<td>Engage</td>
<td>3</td>
<td>3 to 15</td>
<td>7.1 to 12.8</td>
<td>$r = 0.78$</td>
</tr>
</tbody>
</table>
Table 2
Categorization of Student Comments about Learning Objects

### Learning

<table>
<thead>
<tr>
<th>Category Label</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Challenge</td>
<td>Refers to the ease/difficulty of the concepts being covered. Basically whether the content level of the LO matched the student’s cognitive level/understanding. Code “it was easy” in here, but not “it was easy to use.”</td>
</tr>
<tr>
<td>Learn</td>
<td>Student comments about a specific or general learning/teaching issue involved in using the LO</td>
</tr>
<tr>
<td>Visual</td>
<td>The student mention as visual feature of the LO that helped/inhibited their learning</td>
</tr>
</tbody>
</table>

### Engagement

<table>
<thead>
<tr>
<th>Category Label</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compare</td>
<td>Student compares LO to another method of learning</td>
</tr>
<tr>
<td>Engage</td>
<td>Student refers to program as being OR not being fun/enjoyable/engaging/interesting</td>
</tr>
<tr>
<td>Technology</td>
<td>The student mention a technological issue with respect to using the LO</td>
</tr>
</tbody>
</table>

### Quality

<table>
<thead>
<tr>
<th>Category Label</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animate</td>
<td>Refers to quality of animations/moving pictures</td>
</tr>
<tr>
<td>Audio</td>
<td>Refers to some audio/sound aspect of the learning object</td>
</tr>
<tr>
<td>Easy</td>
<td>Refers to clarity of instructions or how easy/hard the LO was to use. It does not refer to how easy/hard the concept was to learn.</td>
</tr>
<tr>
<td>Graphics</td>
<td>Refers to static picture or look of the program (e.g., colours)</td>
</tr>
<tr>
<td>Help</td>
<td>Refers specifically to help/hints/instructions/feedback provided by the LO</td>
</tr>
<tr>
<td>Interactive</td>
<td>Student refers to some interactive part feature of the LO</td>
</tr>
<tr>
<td>Control</td>
<td>Refers to student control of choice/pace in using the LO</td>
</tr>
<tr>
<td>Organization/Design</td>
<td>Refers to quality of organization/design or the LO</td>
</tr>
<tr>
<td>Text</td>
<td>Refers to quality/amount of text in LO</td>
</tr>
<tr>
<td>Theme</td>
<td>Refers to overall/general theme or CONTENT of LO</td>
</tr>
</tbody>
</table>
### Table 3

Varimax Rotated Factor Loadings on Learning Object Evaluation Scale for Students (LOES-S)

<table>
<thead>
<tr>
<th>Scale Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-Learn 1 - Interact</td>
<td>.881</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-Learn 2 - Feedback</td>
<td>.847</td>
<td>.680</td>
<td></td>
</tr>
<tr>
<td>S-Learn 3 - Graphics</td>
<td>.732</td>
<td>.606</td>
<td></td>
</tr>
<tr>
<td>S-Learn 4 – New Concept</td>
<td>.789</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-Learn 5 – Overall</td>
<td>.859</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-Quality 6 - Help</td>
<td>.704</td>
<td>.649</td>
<td></td>
</tr>
<tr>
<td>S-Quality 7 - Instructions</td>
<td></td>
<td>.877</td>
<td></td>
</tr>
<tr>
<td>S-Quality 8 - Easy to use</td>
<td></td>
<td>.873</td>
<td></td>
</tr>
<tr>
<td>S-Quality 9 - Organized</td>
<td></td>
<td></td>
<td>.808</td>
</tr>
<tr>
<td>S-Engagement 10 - Theme</td>
<td></td>
<td></td>
<td>.891</td>
</tr>
<tr>
<td>S-Engagement 11 - Motivating</td>
<td></td>
<td>654</td>
<td>.823</td>
</tr>
<tr>
<td>S-Engagement 12 – Use again</td>
<td></td>
<td>.601</td>
<td>.812</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACTOR</th>
<th>EIGENVALUE</th>
<th>PCT OF VAR</th>
<th>CUM PCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.70</td>
<td>55.8</td>
<td>55.8</td>
</tr>
<tr>
<td>2</td>
<td>1.11</td>
<td>9.3</td>
<td>65.1</td>
</tr>
<tr>
<td>3</td>
<td>0.70</td>
<td>5.9</td>
<td>71.0</td>
</tr>
</tbody>
</table>
Table 4

Correlations Among Learning Object Evaluation Scale for Students (LOES-S) Constructs

<table>
<thead>
<tr>
<th></th>
<th>S-Learn</th>
<th>S-Quality</th>
<th>S-Engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-Learn</td>
<td>1.00</td>
<td>0.68 *</td>
<td>0.74 *</td>
</tr>
<tr>
<td>S-Quality</td>
<td></td>
<td>1.00</td>
<td>-0.64 *</td>
</tr>
<tr>
<td>S-Engagement</td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

* \( p < .001 \) (2-tailed)
Table 5
Correlations Among LOES-S and LOES-T Constructs

<table>
<thead>
<tr>
<th></th>
<th>S-Learn</th>
<th>S-Quality</th>
<th>S-Engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-Learn</td>
<td>0.47 ***</td>
<td>0.47 ***</td>
<td>0.44 ***</td>
</tr>
<tr>
<td>T-Quality</td>
<td>0.45 ***</td>
<td>0.45 ***</td>
<td>0.43 ***</td>
</tr>
<tr>
<td>T-Engagement</td>
<td>0.25 *</td>
<td>0.33 **</td>
<td>0.39 *</td>
</tr>
</tbody>
</table>

*  $p < .05$ (2-tailed)

** $p < .01$ (2-tailed)

*** $p < .001$ (2-tailed)
Appendix A – Learning Object Evaluation Survey - Students

<table>
<thead>
<tr>
<th>Learning</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Working with the learning object helped me learn.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. The feedback from the learning object helped me learn.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. The graphics and animations from the learning object helped me learn.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. The learning object helped teach me a new concept.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. Overall, the learning object helped me learn.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quality</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. The help features in the learning object were useful.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>7. The instructions in the learning object were easy to follow.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>8. The learning object was easy to use.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>9. The learning object was well organized.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Engagement</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>10. I liked the overall theme of the learning object.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>11. I found the learning object motivating.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>12. I would like to use the learning object again.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

13. What, if anything, did you LIKE about the learning object?

14. What, if anything, did you NOT LIKE about the learning object?