The Effectiveness of Discussion Forums in On-line Learning

Barbara A. Lewis

University of South Florida, USA.

Abstracts

Português

O presente estudo investiga a efetividade da aprendizagem no contexto de cursos online, utilizando duas formas alternativas de atividades práticas: foruns de discussão assíncrona online e exercícios/testes (quizzes) resolvidos individualmente. O estudo foi desenvolvido em cursos regulares existentes onde a efetividade da aprendizagem era formalmente avaliada, através de testes objetivos oriundos do conteúdo específico do curso. A relativa efetividade das atividades práticas pode ser justificada em direções diretamente opostas se alguém argumenta sob os pontos de vista das teorias objetivista ou construtivista. Os resultados obtidos neste estudo, em parte, parecem apoiar a posição objetivista; por outro lado parecem apoiar a posição construtivista. Uma análise posterior dos dados coletados sugere que estes dois resultados, aparentemente contraditórios, podem, de fato, ser compatíveis, quando é considerado o fator de intensidade do envolvimento com as atividades das discussões colaborativas. A conclusão tentativa é de que as atividades de discussão de grupos online, devem alcançar certo nível de intensidade e compromisso por parte dos participantes, para que resulte numa efetiva aprendizagem.

Inglês

This study investigated the learning effectiveness in online course contexts of two alternative forms of practice activities: asynchronous online discussion forums and individually completed quizzes. The study was conducted in existing regular courses, where learning effectiveness is formally assessed by means of objective tests derived from the subject matter content of the course. The relative effectiveness of the two forms of practice activities may be supported in directly opposed directions if one argues from constructivist or objectivist theoretical positions. The results obtained in this study in part seemed to support the objectivist position and, in another part, seemed to rather support the constructivist position. Further analysis of the data collected seems to suggest that these two apparently contradictory results may in fact be compatible when the factor of the intensity of engagement in collaborative discussion activities is taken into consideration. The tentative conclusion is that online group discussion activities must reach a certain level of intensity and engagement by the participants in order to result in effective learning.

Espanhol

Este estudio investigó la eficiencia del aprendizaje en contextos de cursos online de dos formas alternativas de actividades prácticas: foros de discusión asincrónica online y pruebas cortas (quizzes) completadas individualmente. El estudio conducido se realizó en cursos regulares en los cuales la eficiencia del aprendizaje es evaluada formalmente por medio de pruebas objetivas, originadas del contenido de las materias del curso. La efectividad relativa de las dos formas de actividades prácticas puede ser justificada en direcciones directamente opuestas si uno argumenta desde posiciones teóricas constructivistas u objetivistas. Los resultados obtenidos en este estudio en parte parecen apoyar la posición objetivista y, por otro lado, parecen apoyar la
posición construtivista. Analisis posterior de los datos colectados indican que estos dos resultados, aparentemente contradictorios, pueden, de hecho, ser compatibles cuando es considerado el factor de intensidad del compromiso con las actividades de discusión colaborativa. La conclusión tentativa es que la actividad de discusión de grupos online debe alcanzar cierto nivel de intensidad y compromiso de los participantes para que resulte en aprendizaje eficiente.

Introduction

Do online learning environments (web courses) work? Do people learn in these environments? These are complex questions, which are as yet little investigated through systematic research. The literature on the topic is large and growing, but most of it is anecdotal rather than empirical. The many outstanding research questions will not be resolved quickly, since many variables need to be accounted for and control groups established for comparisons, which is a difficult task in real-life "intact" educational environments (Mayadas, F., 1997).

Early studies of online education focused on the viability of online instruction when compared to the traditional classroom. Recently, researchers have begun to examine instructional variables in courses taught on-line. Berge (1997) conducted a study of forty-two post-secondary online instructors to discover strategies that educators might use to improve their online teaching. The instructors indicated that they believed learner-centered strategies to be more effective than instructor-centered strategies. They also indicated that they preferred the following methods: discussion, collaborative learning activities, and authentic learning activities. However, what was not discussed in the study was the effect the strategies had on the students. The purpose of the study reported here is to investigate the learning effects of one of these "preferred" strategies: on-line discussions.

Background and Rationale for the Study

In recent years, partially as a result of the so-called "technology revolution" and partially due to paradigmatic shifts in educational philosophy, both the theories and the practice of instruction have undergone significant change. In the area of learning theories, there has been a shift from a behaviorist to a constructivist view of learning as a process involving the construction of knowledge. This, in turn, has led to an increasing emphasis on collaborative learning strategies, in which people work together in small groups. The physical environment of learning is also shifting ever more from face-to-face classroom instruction, to distance-learning on the Internet.

The substitution of interactive "CAI" tutorial sequences, or individually completed quizzes, by online group discussions is observed to be an increasingly common practice among teachers who modify previously existing courses for online delivery. This trend is often justified from the standpoint of Collaborative Group Learning principles drawn from theories of Active Learning based on modern educational philosophies such as Constructivism. However, the available research data that would confirm these claims is scarce and inconclusive. Furthermore, given that the popularity of this trend seems to have grown with the increasing availability of efficient technology for the organization and management of threaded discussions, one may question whether theoretical principles or technological fashion are the real driving forces.
It also seems that some of the specific new strategies that are being implemented in the name of new theoretical positions do not always exhibit the characteristics that these strategies should (theoretically speaking) embody. In some cases it seems that the changes are driven more by the appearance and availability of the new technologies than by any coherent set of theoretical principles. The goal of this study was to investigate the extent to which one specific change in methods and media, namely the use of asynchronous discussion environments as a component of online courses can be seen to be theory-driven or technology-driven. Another motivation for the study arose from the desire to understand the effectiveness of such discussion forums on students' achievement scores. Among the many as yet unanswered questions regarding Web-based courses is whether the use of asynchronous online discussion activities, as a means for providing opportunities for practice and learning, is necessarily an improvement over previously used strategies, such as quizzes.

The theory and practice of the discipline of instructional design also suggests that in order to implement a new instructional design, based on a different theory of learning, it is usually necessary to modify not only one, but maybe all or most of the components of a lesson (Romiszowski and Chang, 2001; Dills and Romiszowski, 1997). However, it is currently quite common to utilize the newly available online discussion environments as the "practice" component of lessons that are otherwise unaltered in their basic instructional design. Existing content-presentation materials, previously used in conventional courses, often espousing other learning theories, are posted to the Web without any modification. The same final evaluation tests and procedures are employed, regardless of the implied modifications to the underlying course philosophy and shift in key objectives from the content to the process of learning. The present study has intentionally selected just such a context for its investigation.

The course selected for the study

An existing course that has for some time been offered as a conventional face-to-face course is now also being offered as an online course. This course is based on a well-established basic textbook that not only is a major source for the course content, but also includes a large questions bank from which instructors may create a variety of more-or-less equivalent learning assessment instruments and practice quizzes. In the process of transforming the conventional course to an online version, little instructional design change was introduced as regards the "presentation" phase, in that the same textbook was made available online and similar instructor advice and support was offered. Also, little change occurred with respect to the final "test" or "assessment" phase, in that the same questions bank was used to generate final examinations. However, some of the instructors involved chose to modify the "practice" phase by introducing online discussion activities in place of the previously used quizzes. The present study was designed to investigate the learning effectiveness of this change in the basic instructional design.

This particular course is a fifteen-week on-line course in a major university setting. The course and the instructional materials it uses (i.e., the content of twelve chapters of the set book, the test bank and any tests and unit quizzes derived from the bank) is a standard on-line course that is offered by 3 different instructors each semester at the university. The enrollment is 50 students per course. Therefore, on an average, 150 students per semester take the on-line version of the course, using the same course materials. The entire course syllabus, quizzes, and discussion activities are available on-line in a WebCT course shell.

Experimental procedure
An intact cohort of fifty students, registered to take the abovementioned course was randomly sub-divided into two experimental groups who were subjected to different treatments as regards the "practice" phases of the online lessons that compose the course. All students participated in quizzes for some of the lessons and in online discussions for other lessons, according to the experimental design explained below. This procedure allowed the investigator to compare the learning effectiveness of the two alternative practice procedures and also to investigate some other secondary questions. The following procedures were applied to the assignment of the participants to the treatment sequences and measurement of the results. Each participant:

- completed an online pretest which was based upon the information contained in 12 chapters of the required textbook;
- read the book and the lecture notes, one chapter per course unit;
- completed six online quizzes for six of the course units (based on randomized assignment to one of two groups: Group 1 in odd and Group 2 in even units);
- completed six threaded discussion forums for the other six course units, which were based on questions posted by the instructor on issues in the unit.
- completed an online posttest based upon information in the textbook (exactly the same assessment procedure that has been used for years for grading both on-line and face-to-face versions of the course);
- completed an end of course evaluation questionnaire.

The tests were taken from the test bank prepared by the publisher of the book used in the course. This book and test bank have been used for the past three years at the university. As stated above, the course is offered three times a semester as an on-line course for a total of nine times a year. Besides the on-line version of the course, this course is also offered three times a semester as a traditional course using the same test bank. Therefore, even though there is no available statistical analysis of the reliability of the test items, it could be inferred that the test questions do have general acceptance by expert teachers of the subject as a valid instrument by which to measure learning of the course material. Different versions of the assessment instrument (i.e., test) have been used at least six times a semester (including traditional and on-line courses), three times a year, over a period of three years, for a total of fifty-four times. However, the most important point to make as regards the present study is that the statistical comparisons of different groups undergoing different treatments are all based on the results obtained by participants on just one unique version of the test - the post-test that was taken by all the participants in the cohort selected for the study.

Overall Results

Fifty students began the class; however, only 37 students finished the course. Thirteen students either dropped out of the course or took an incomplete in the course. The remaining 37 students remained in the same random groups and subgroups as assigned in the beginning of the course. The first step of the experiment involved the administering of a pre-test. The overall pre-test scores were rather variable and in some cases quite low. At the end of the semester, an equivalent post-test was administered. It is apparent that students did improve from the pretest to the posttest. The overall mean score of the pretest was M=27.89 and the overall mean score of the posttest was M=625.30.

However, the only reason for administering a pre-test was to verify that the randomly selected groups were indeed equivalent. In the event, Group 1 pretest scores were a little higher than Group 2 pretest scores, but this difference was not found to be statistically significant. Once this was established, all comparisons between the groups
were made on the basis of post-test scores. Each posttest score was divided into the 12 chapter units scores. Each unit consisted of 5 questions worth 12 points each. It is apparent from the bar graph (Figure 1) that there are differences among the subunit scores. In fact, when Group 1 scored higher than Group 2 the difference was extremely notable to the eye. However, when Group 2 scored higher than Group 1 the differences were not as notable, except in the one case of unit 11 where it is extremely notable that Group 2 did better than Group 1.

![Sub Unit Posttest Mean Scores](image)

An analysis was conducted on the subunit scores to check for significance of this rather surprising result. Several one-way ANOVAs were performed to test the null hypothesis: "there is no difference in the learning outcome for those who engage in discussion activities versus those who complete the quizzes". This analysis revealed that the null hypothesis is accepted for subunits 1, 3, 5, 6, 7, and 9. However, the null hypothesis was rejected for subunits 2, 4, 6, 8, 10, 11, and 12.

This finding is interesting in that the Chapters 2, 4, 8, 10, and 12 are the chapters for which Group 2 did the discussion forums and Group 1 did the quizzes. These results, taken on their own, seem to suggest quite strongly that the quiz-taking activity generally leads to superior post-test performance than the discussion activity. This indeed was the result we expected to find, from an instructional-design-driven vantage point, on the grounds that the quizzes were identical in nature to the post-test. Both were practicing and testing comprehension of the basic content of the chapters, whereas the discussion activities could be expected to focus on other categories of learning that were not being measured by the type of post-test that was used. This was the instructional design "flaw" in much of current on line course design practice that we were hoping to reveal.

However, the other "half" of our results did not tally with this finding. The only time when there was significance when Group 2 did the quizzes and Group 1 did the discussion forums was in subunit 11. In all the other 5 such cases, the differences were not significant. Furthermore, this one case of significance may possibly be explained by the drop out of students towards the end of the course, which happened to be
particularly concentrated in Group 1. The question that arises out of the data, therefore, is why is there generally no significance when Group 2 takes the quizzes and Group 1 engages in on-line discussion.

To explore this question, we looked at the content of the online discussions. We reviewed the number of messages read and number of messages posted to see if any differences may have had an effect on the posttest scores. Could the significance be caused by a correlation between the number of messages read, number of messages posted and the posttest scores? To further investigate the difference between the mean scores of Group 1 and Group 2 on the subunits, we looked at the correlation between the groups and the number of messages posted to the threaded discussion and the number of messages read. A one-way ANOVA was conducted on both the messages read by the students and messages posted by the student. There was a significant difference on messages read by students between groups. However, there was no significant difference on messages posted within the groups.

This finding suggests, however tenuously, that the varying amount of effort and frequency of participation of the members of one group, as compared to the other group, led that group to get "more value" out of the discussion activities and thus compensate for the "handicap" imposed by the lack of a practice exercise that was directly aligned to the final evaluation instrument. In order to check whether this agreed with the student's own perceptions of the experience, a specially designed course evaluation questionnaire was administered at the end of the course.

Discussion of the Results

Constructivist theory states that students should be encouraged to construct their own knowledge. Computer-Mediated-Communication, it is argued, effectively supports constructivism because of the emphasis on access to resources and the extent of collaboration between students promoted through the use of discussion boards. Therefore, many constructivists argue, students in an online environment can construct their knowledge through active learning and collaboration and, therefore, would presumably learn more effectively.

Engagement theory suggests that learners must be actively engaged in meaningful tasks for effective learning to take place (Kearsley and Schneiderman, 1998) and one means of providing such meaningful tasks is to engage the students in discussions. The students in the present experiment were actively engaged in threaded discussion activities at several points throughout their course of study. These discussion boards created an environment where the students actively read the comments of their peers and "conversed" by exchanging messages with other students and the instructor.

Both engagement theory and collaborative learning theory suggest that the use of discussion forums brings the students directly into contact with the content material of the course instead of leaving them on the outside as passive learners. Through this interaction, it is postulated, students are building their knowledge instead of relying on simple memorization skills. If these theoretical positions are valid, one could expect the use of discussion forums to be more effective than quizzes as a means of promoting learning.

Researchers also argue that collaborative learning and social interaction play a major role in cognitive development. Collaborative learning is the "acquisition of knowledge, skills or attitudes that take place as a result of people working together to create meaning, explore a topic or improve skills" (Graham and Scarborough, 1999). Hiltz
(1997) states that collaborative learning is crucial to the effectiveness of online learning environments. However, both these authors seem to espouse online learning mainly because it offers tools for collaboration and so is "in tune" with the authors' theoretical and philosophical views on education in general and the learning process in particular. We see a certain circularity in the arguments presented in the literature. This lack of clarity in the arguments makes it particularly important to investigate the relative effectiveness of the two "levels of interaction" represented by the two most-used forms of online learning exercises: individual quizzes and group discussion forums.

The significant differences found in the present study in subunits 2, 4, 8, 10 and 12 would have the author claim that there is a difference between those who engaged in discussion activities and those who engaged in quizzes on the posttest. However, the significance is in a negative direction with respect to the discussion activities. The significant differences showed Group 1 posttest scores were in most cases higher in the units where Group 1 students engaged in individual quizzes rather than online discussions. On the other hand, the remaining subunits did not show any significant differences. Most of these sub-units are those when Group 2 studied by means of quizzes and Group 1 by means of discussion forums. This subset of results seems to support the null hypothesis - that there is no difference in learning under the two conditions. The results of the study seem, therefore, to be somewhat inconclusive, but inconclusive in a systematic manner. There seems to be a pattern in the results that may possibly be interpreted by further analysis.

It is apparent from the distribution of mean scores of the subunit posttest scores that the discussion forums had a somewhat different effect on the posttest scores of Group 1 than they had on Group 2. It seems that Group 1 students who participated in the discussion activities scored just as well for these subunits as the Group 2 students who took the quizzes on those subunits. These partial findings do not support the theoretical positions discussed above. In the first partial result, the Group 1 superiority when studying by means of quizzes is diametrically opposite to the result predicted by the theories we have discussed. It is rather more in line with the positions of traditional pedagogy that tend to support the use of practice activities that are directly equivalent to the final evaluation activities. It supports the "objectivist" approach that tends to design the in-course learning activities on the basis of an analysis of the end-of-course performance objectives. However, the second partial result is not so conclusive. Although the discussion groups do no better, they also do no worse than the groups that completed quizzes.

Let us examine these findings from yet another theoretical position - the "objectivist" theory of instructional design. This position has a long history of practical use and acceptance. It is rather incorrect and unfair to label the position as "behaviorist", because it really represents the established practice of the teaching profession from times way before the development of behaviorism. However, this position did tend to get formalized as a result of the growing popularity of the use of behavioral objectives as a basis for the design of learning activities. The practical influence of "programmed instruction" models reinforced the widespread acceptance, almost as an axiom, of the principle of designing the learning activities as a "mirror image" of the final evaluation activities. In the case of this particular study, the objectivist position would argue that we should expect the quizzes to be more effective learning activities than the discussions, because they better reflect the final test conditions used to evaluate the learning. Once more, however, one must observe that, in the present study, one part of the results supports this position, but the other part does not.
Further light is, however, shed on the results of this study if one examines the objectivist position a bit more critically. The partial result that students who participated in the discussion activities scored just as well as those who took the quizzes is in line with Mouton's (1988) findings that success on lower level testing can be achieved by the review of "higher-order learning" problem solving questions during the practice assignments. In his study, Mouton looked at what types or combination of types of practice activities should be provided to students, studying through mediated self-instruction. The finding of the study showed that a "more stable and durable memory trace results if deeper cognitive processing occurs during encoding." (p. 97) and "students when engaged in higher level thinking questions will do as well on lower level thinking test items as students just doing lower level thinking questions".

Also pre-dating the constructivist movements of today, Bloom (1981) suggested that, in order to be independent and active learners, the learners should engage in so-called "higher-level thinking". They should also "possess the ability to learn and solve problems, be intrinsically motivated, and possess a degree of social responsibility to interact with others in the acquisition of learning". Using the logic of Mouton and Bloom, the use of online discussion forums can be postulated to serve as an avenue for learners to obtain higher levels of achievement, even on lower-level rote-memory test instruments, than by means of participation in lower-level forms of learning activities, such as quizzes. From this theoretical position, the use of higher level thinking questions and discussions does not hinder but enhances a student's learning, even if tested by lower level thinking tests. This theoretical analysis helps to explain the partial finding in the present study that Group 1 students studying in the "higher-order-thinking" mode of the discussion forum did just as well as Group 2 students who studied these same subunits in the "lower-order-thinking" mode that was a "mirror image" of the final test conditions.

However, we still have the other partial result that seems to support the conventional objectivist position of designing the learning activities as a mirror-image of the testing procedures. It is difficult to escape the conclusion that, despite the apparent equivalence of the two groups, as demonstrated by means of analysis of overall pre-test and post-test scores, something differentiated them during the course of the study. One factor that may have played a part is the intensity and frequency of participation in the group discussions. As shown below, this seems to have had an influence on final performance.

Students who participated frequently and intensively in the online discussions could be expected to have benefited from the higher level thinking activity more than those students who engaged less thoroughly and less frequently in the discussions. If it can be shown that Group 1 students tended to fall into the first category and Group 2 students into the second category, we may have an explanation that fits both of the partial results of this study. If the Group 2 students did not participate as frequently and as deeply as they should have in the discussion forums, then they did not really engage in "deep" intellectual learning activities and therefore did not do as well as when they used objective quizzes as the practice component of their study. On the other hand, if the Group 1 students did participate more thoroughly in the discussion activities, and therefore did engage in "deep processing" of the course content, then as suggested by Bloom, Mouton and others, their performance on any form of lower-level test would also be enhanced and so, in effect they may do better, or at least as well, as the Group 2 students who had the apparent benefit of practice activities mirrored on the final tests.
Did the two groups participate differently in the course? To further research this question, a closer examination of group activity was conducted. As noted earlier, students in each group were required to complete six threaded discussion assignments for alternating chapters. The students were expected to read and respond to the instructor's question for each threaded discussion and other students' responses. The lowest number of threaded discussion messages read by students was 17 messages and the highest number of messages read by students was 124 messages. The overall mean of threaded discussion messages read by students was 52 messages. Students in Group 1 read more messages then students in Group 2 (Group 1, M = 61 messages and Group 2, M = 43 messages).

Palloff and Pratt (1999) claimed that interaction and collaboration become critical in Web-based training. They also suggested that the successful online learner is a "noisy learner" who is active and creative in the instructional environment. Students in Group 1 were more active than students in Group 2. This is apparent from the number of messages read by the students. A review of the number of messages read and number of messages posted was conducted to see if any differences may have had an effect on the posttest scores. Could the significance be caused by a correlation between the number of messages read, number of messages posted and the posttest scores? This analysis showed that there was a significant difference on messages read by students between groups. However, there was no significant difference on messages posted within the groups.

Conclusions

As already noted above, there were significant differences in subunits 2, 4, 8, 10 and 12 that would have the author reject the "null hypothesis" and claim that there is a difference between those who engaged in discussion activities and those who engaged in quizzes on the posttest. However, the significant difference showed Group 1 posttest scores were higher in the units where students completed the quizzes. This suggests that the relatively low levels of participation in discussion forums for these same subunits by Group 2 students were insufficient for these students to really "deep process" the content of these subunits. As a result, the rote-learning activities of Group 1 led to superior test scores, not so much because Group 1 students learnt more from participating in the quizzes, but rather because the Group 2 students learnt less, as a result of inadequate participation in the discussions.

However, in the other "half" of our results, in 5 (out of 6) cases, the differences were not significant. Furthermore, this one case of significance may possibly be explained by the drop out of students towards the end of the course, which happened to be particularly concentrated in Group 1. The question that arises out of the data, therefore, is why is there generally no significance when Group 2 takes the quizzes and Group 1 engages in on-line discussion. One possible answer is that the significantly higher frequency of reading and posting of messages to the discussion forums by Group 1 students led these students to engage in "deep processing" of the content in a manner that the Group 2 students had in general not achieved. As demonstrated by Mouton (1988), when students participate effectively in higher order learning activities (in the present case, in discussion forums), they may do just as well on lower order learning assessment as those who did the lower order learning practice activities.

Thus, a possible, though by no means proven, interpretation of the results of this study is that the difference between Group 1 and Group 2 scores is due to the varying amount of effort and frequency of participation in group discussion activities. The higher level of engagement of Group 1, as compared to Group 2, led that group to get
"more value" out of the discussion activities and thus compensate for the "handicap" imposed by the lack of a practice exercise that was directly "mirrored" on the final evaluation.

Further research will be required in order to establish whether this new hypothesis is consistently supported in practice. If it proves to be supported, one may gain some important insights into the factors that must be designed into online learning activities in order to ensure that they are effective learning experiences as measured and evaluated by the conventional, content-based, criteria that are commonly utilized by most educational systems. Finally, we may add that the present study once more illustrates the importance of adopting a theory-and-research-based instructional design approach to Web-based education and training.

References


