ASSESSMENT AND COLLABORATION IN ONLINE LEARNING

Karen Swan
Research Center for Educational Technology, Kent State University

Jia Shen
School of Management, New York Institute of Technology

Starr Roxanne Hiltz
New Jersey Institute of Technology

ABSTRACT
Assessment can be seen as the engine that drives student course activity, online or off. It is particularly important in encouraging and shaping collaborative activity online. This paper discusses three sorts of online collaborative activity—collaborative discussion, small group collaboration, and collaborative exams. In each of these areas, it provides both theoretical grounding and practical advice for assessing, and so encouraging, collaboration in online courses.

KEYWORDS
Online Collaboration, Assessment, Course Discussion, Small Group Work, Collaborative Exams

I. INTRODUCTION

The theme of this paper is the importance of assessment to learning; that what is assessed is what is valued, and if you value collaboration as an instructor, you need to find ways to motivate students and to assess collaborative activity. However, in most online courses, traditional instructor-centered examination remains the primary means for assessing student performance, and collaborative learning is undervalued and so marginalized. This paper addresses that issue. The paper begins with an overview of the changing nature of assessment in education and addresses assessing online collaboration—why it is important and why it is difficult. It then explores in greater depth two common and quite different kinds of online collaboration, collaborative online discussion, and collaborative small group projects. Finally, findings from a recent dissertation which introduced and measured the effectiveness of innovative collaborative examination procedures are summarized. In each or these areas we lay out the issues, theoretical and practical involved in assessing online collaboration, as well as give examples of how to assess (and so encourage) individual and group work for differing sorts of activities.

A. What is Assessment?

Value in any instructional system comes from assessment; what is assessed in a course or a program is what is valued; what is valued becomes the focus of activity. The link to learning is direct. Instructors signal what knowledge skills and behaviors they believe are most important by assessing them. Students quickly respond by focusing their learning accordingly.

While assessment is often equated with tests, exams, and evaluations, the term assessment can be used more broadly, as shown by the following two definitions provided by educational researchers:

> *Assessment is defined as the systematic basis for making inferences about the learning and development of students. More specifically, assessment is the process of defining, selecting, designing, collecting, analyzing, interpreting, and using information to increase students' learning and development.* [1]

> *Assessment is an ongoing process aimed at understanding and improving student learning. It involves making our expectations explicit and public; setting appropriate criteria and high standards for learning quality; systematically gathering, analyzing, and interpreting evidence to determine how well performance matches those expectations and standards; and using the resulting information to document, explain, and improve performance.* [2]

Critics of objective exams and standardized tests assert that such assessments measure surface learning, examining only at the level of unrelated bits of memorized information [3], when course objectives usually involve deeper levels of learning. There are many ways to define a higher level of "deep" learning. For, example Bloom's [4] well known taxonomy of cognitive process levels begins with knowledge and comprehension and moves up to higher levels such as application analysis, synthesis and evaluation. With the Total Quality Movement (TQM) and its adoption in education [5, 6], assessment has evolved from measuring students' learning to an important part of the education process itself [7]. TQM advocates identifying the critical processes in an educational enterprise and assessing them as well as their products. Learner-centered assessment processes [8] have been proposed to shift the focus from instructors and teaching to students and the learning process itself. The authors encourage the development and use of artifacts such as projects, papers, performances, portfolios, or exhibitions that evaluate higher-order thinking. Such performance-based assessment techniques require students to directly reveal the abilities and skills and knowledge that professors desire to develop in their students.

B. Assessing Collaborative Learning in Online Environments

Many theoretical and empirical analyses emphasize the importance of active participation and collaboration among students in promoting the effectiveness of online learning. However, in most online courses, traditional instructor-centered examination remains the primary means for assessing student performance, and collaborative learning is undervalued and so marginalized. In a large part, this is because the assessment of collaboration requires a radical rethinking of assessment methodologies. Three issues are involved: the variety and kinds of goals for online collaboration, the complexity of assessing both individual and group behaviors, and collaboration on assessment itself.

The first issue is how various are the kinds and learning goals of online collaboration and so how difficult it is to address the assessment of collaboration generally. Some examples of the diversity of focus among collaborative activities in online environments are the collaborative construction of knowledge bases [9, 10], the collaborative investigation of scientific phenomenon [11, 12], group engagement in game-like learning tasks [13] or simulations [14], peer review and evaluation of learning products [9], online peer mentoring [15], collaborative analysis of case studies [16], and collaborative discussion groups [17, 18]. Even within these various groupings, one single sort of assessment will not be appropriate because learning goals vary from implementation to implementation. For example, Nachmias, Mioduser, Oren & Ram [19] distinguish between structured and emergent collaboration schemes. In the latter sorts of collaboration activities, assessment must also emerge.

What is consistent across the varieties of online collaboration is that collaborative learning will be more successful when it is valued, and that any such assessment should begin with a very specific understanding of desired learning. For example, in some collaboration activities, learning to collaborate is seen as an important part of what is to be learned; in others, it is merely a means to an end. In some collaborative activities, collaboration is focused on producing a group project, in others it is designed to improve the quality of individual work. Specific requirements for collaboration, including detailed assessment rubrics focused on critical collaborative processes, will help students achieve desired goals.

The second issue is that collaboration is a complex activity which involves both individual and group effort. To encourage collaboration, both aspects must be assessed. Johnson and Johnson [20, 21], for example, contend that the key to successful cooperative learning is maintaining both individual accountability, in which students are held responsible for their own learning, and positive interdependence, in which students reach their goals if and only if the other students in the learning group also reach theirs. The way to ensure individual accountability and positive interdependence, according to Johnson and Johnson, is to assess both individual and group learning.

A simple example of this kind of assessment using summative testing is to give each student a grade based on some combination of their test score and the average score for their group. Another frequently used scheme is to give a common assessment for a group project and have group members rate their peers' contributions which are then averaged for individual grades. Unfortunately, these kinds of grading protocols are not often seen in online courses where the common approach is to assess either individual effort e.g., (online discussion participation) or group products (collaborative projects).

The third issue is the role of collaborative assessment. Some argue that if collaboration is an essential feature of successful online learning, then assessments as well as activities should be collaboratively designed. Some recent procedures have been described that incorporate student active participation and collaboration into the assessment process itself. Participation and collaboration have been integrated into various phases of collaborative assessment, such as collaborative development of the grading scheme [22], collaborative question composition [23], collaborative question answering [24], collaborative examinations [25], and peer and self-grading [26, 27].

II. ASSESSING & ENCOURAGING COLLABORATIVE DISCUSSION

Online discussion has been an object of interest to researchers for at least two decades because of the potential it holds to support learning. Many researchers note that students perceive online discussion as more equitable and more democratic than traditional classroom discussions because it gives equal voice to all participants [28, 29]. Online asynchronous discussion also affords participants the opportunity to reflect on their classmates' contributions while creating their own, and to reflect on their own writing before posting it. This creates a certain mindfulness and reflection among students [30, 31, 32]. In addition, many researchers have noted the way participants in online discussion perceive the social presence of their colleagues, creating feelings of community [33, 34, 35, 36, 37]. Indeed, an increasing number of studies have examined the perception of interpersonal connections with virtual others as an important factor in the success of online learning [38, 39, 40, 41, 42]. Such findings have led educators to conclude that asynchronous online discussion is a particularly rich vehicle for supporting collaborative learning.

Researchers have also found that successful online collaborative discussion is directly linked to its assessment [43, 44, 45, 46]. Simply put, this means that to encourage collaborative discussion one must

grade it. Discussion participation must count for a significant portion of the course grade and individual discussion postings must be individually assessed. A requirement of a particular number of discussion postings per week or per course module will help ensure students participate in discussion.

Assessment can be done by counting things like the number, regularity, and length of contributions. The problem with this approach is that if students know this is the basis of grading, they may simply load the class discussion with items that are not very thoughtful or original, or perhaps not even on the subject. For example, many instructors use tools that automatically count the number of messages written by a student as a proxy for that student's participation [47]. However, care should be taken to avoid counting superficial posts [48]. One way to detect these 'low value' messages is to judge the reaction of other students. If students are not responding to messages written by a certain student, then either the postings are of little value, or they are too verbose and the other students are not taking the time to read and respond to those messages. Either way, responses to web conference messages can act as a proxy for the value of the student interaction.

There are several ways of providing visualizations of response patterns to aid assessment of the extent to which a student's postings are stimulating responses from others. One is to create bar graphs which show the number of postings made and number of responses received by all of the students. Another is to automatically create student-centered "who to whom" social network diagrams [49]. By visualizing and comparing a student's social network through their social graphs, one can potentially gain a better understanding of a student's involvement within a class.

Another potentially more effective approach is to develop grading rubrics that assess specific discussion behaviors. While grading is the most common form of assessing students' learning, rubric scoring provides for a more finely-detailed characterization of students' behaviors than simple grading. Rubrics typically consist of: (1) a set of categories—features or aspects of student work that are of interest, such as "use of course concept x" or "degree of reflection"; and (2) hierarchical levels of performance within each category, such as "0—course concept x not used, 1—course concept x inappropriately used, 2—course concept x appropriately used but not justified, 3—course concept x appropriately used and justified." For example, the following rubric (Figure 1) was developed by one of the authors. It should be noted that her goal for course discussion was for students to link course concepts to their own experience and thus the rubric assigns up to two points for messages that meet those criteria.

	0 points	1/2 point	1 point
Links to course material	No links to course material	Unelaborated or poorly elaborated links to course material	Links to course material with elaboration and/or additional information
Links to personal experience	No links to personal experience	Irrelevant or unjustified links to personal experience	Relevant and well justified links to personal experience

Figure 1: Rubric for Assessing Discussion in an Educational Computing Course

Despite the effectiveness of rubrics in assessing students' learning, developing an adequate rubric for a given course discussion requires time and, often, multiple revisions. The first consideration in developing a discussion grading rubric is establishing the goal or goals of the discussion. For example, some instructors want to use online discussion to help students learn argumentation techniques [50], whereas others might be more interested in students coming to consensus on a topic. Clearly different assessment rubrics would be needed to encourage each. Thus, the second step in developing a grading rubric is to

identify characteristics of messages that would support the established goal. For example, a discussion rubric aimed at encouraging successful argument might identify such things as clear statement of position, identification of points of agreement and disagreement with previous postings, logical arguments, and so on as characteristics to be evaluated. The third and final step in rubric creation involves then taking each characteristic and specifying differing levels of performance for each and assigning scores for these (see Figure 1).

Another way of going about creating a rubric for assessing discussion postings is illustrated in the first part of Figure 2. In this example, the instructor, Pelz [51], identified the goal of the discussion as knowledge construction; that is, he viewed discussion as central to content learning. Thus, he was particularly interested in encouraging student contributions to the discussion that were accurate, original, relevant, and that added to content learning. Instead of creating a rubric that separately specified differing performance levels for each of these characteristics, however, he grouped them together and specified differing performance levels for that group.

Discussion Rubric:
Each discussion post is graded according to the following rubric.

Points	Interpretation	
4	Excellent (A)	The comment is accurate, original, relevant, teaches us something new, and well written. Four point comments add substantial teaching presence to a course and stimulate additional thought about the issue under discussion
3	Above Average (B)	The comment lacks at least one of the above qualities, but is above average in quality. A three point comment makes a significant contribution to our understanding of the issue being discussed.
2	Average (C)	The comment lacks two or three of the required qualities. Comments which are based on personal opinion or personal experience often fall within this category.
1	Minimal (D)	The comment presents little or no new information. However, one point comments may provide important social presence and contribute to a collegial atmosphere.
0	Unacceptable (F)	The comment adds no value to the discussion.
No penalty	Excellent Subject	The subject field contains the main point of the comment. The reader clearly understands the main point of the comment before reading it.
1 point penalty	Minimal Subject	The subject field provides key word(s) only. The reader knows the general area that the comment deals with.
2 point penalty	Subject Field is Unacceptable	The subject field provides little or no information about the comment.

Figure 2: Discussion Rubric for Psychology courses (Pelz, 2004)

What is also interesting about this rubric is that in its second half, points are taken off for inadequate subject headings. Dr. Pelz relates that poor subject headings were making discussions in his online psychology classes hard to follow; hence he adopted this technique to encourage better ones. Thus, another approach to creating discussion rubrics might include identifying discussion behaviors to discourage poor subject headings and introducing disincentives for these.

While the development of rubrics as described above will help improve individual discussion performance, they may not ensure collaborative performance. Recent research by Chia-Huan Ho [52] linking ratings on a rubric derived from Grice's cooperative principles for effective face-to-face discourse [53] to both the numbers of responses generated by an individual posting and the average number of responses generated by particular students suggest that this rubric (Figure 3) may elicit collaboration. It is certainly worth further investigation. In addition, Ho found that students' overall Gricean ratings were also linked to their final course grades, suggesting the value of collaboration.

PT	QUANTITY	QUALITY	RELEVANCE	MANNER
3	The amount of information is sufficient to clearly establish the purpose of the posting.	The posting is a *new contribution* (e.g., novelty, originality), reflective of the student's opinions, ***AND*** is supported by *accurate* evidence/examples.	The posting is on the same topic as both the conference, ***AND*** the previous posting.	The posting is logically organized and has no spelling, punctuation, or grammatical errors; meaning of the posting is clearly presented.
2	There is *slightly* too much or too little information; however, the purpose of the posting is still reasonably clear.	(a) The posting is a *new contribution* that reflects the student's opinions; however, evidence/examples are <u>not</u> provided to support claims; ***OR*** (b) The posting reflects the student's opinions and *accurate* evidence/ examples are provided.	The posting is on the same topic as the conference, but *not* the previous posting.	The posting is adequately organized; if any errors are found, they are so minor that the meaning is still reasonably clear.
1	There is too much or too little information, such that the purpose of the posting is occasionally obscured.	(a) The posting is representative of the student's opinions, yet evidence/examples are <u>not</u> provided to support claims; ***OR*** (b) The posting is largely a re-statement of prior postings ***BUT*** incorporates a *minor new contribution.*	The posting is on the same topic as any of the previous postings, but *not* the conference.	The technical aspect of the posting (e.g., organization, spelling, grammar) has several problems, such that the meaning is occasionally obscured.
0	There is so much or so little information that the purpose of the posting is not understood.	(a) The main idea in the posting is a re-statement of prior postings and *no new contribution* is present; ***OR*** (b) *Inaccurate* evidence/ examples are provided.	The posting is irrelevant to both the conference topic, ***AND*** previous postings.	The posting is poorly organized and/or it has serious errors in sentence structure or usage, thus the posting is hard to understand.

Figure 3: Grice's Cooperative Principles Rating Scale

Another way to provoke collaboration is to develop rubrics that reward collaboration. Rubrics that reward collaboration must focus on discussion responses. They might, for example, only credit responses that cite and either extend or refute previous postings. Another possibility is to assess postings based on the discussion threads they engender, making thread initiators responsible for sustaining collaborative discussion. Pelz [51], for example, makes students responsible for developing questions for each discussion and then grades them, not only on the quality of their questions, but also on the quality and length of the discussion they elicit. As all students are responsible for posting questions to every discussion, this encourages students to support each other by contributing to the various discussion threads. That students also get credit for their individual postings makes them more eager to collaborate A third possibility that is perhaps more applicable to small group (as opposed to whole class) discussions is to grade individual contributions but to also give a group grade based on some average of these. This insures both individual accountability and group interdependence.

Finally, an important means for assessing and encouraging collaborative discussion is to have some sort of outcome or product of discussions which is graded. These kinds of options are again more applicable to small groups because it is difficult to collaborate on a project in a large group. Group members might be asked to collaborate on a discussion summary or to develop collaboratively a case analysis or a solution to a problem. Individuals within the group could then be graded both on their individual contributions using rubrics as discussed above and on the group product. Assessment of group products will be discussed further in the section which follows.

III. ASSESSING & ENCOURAGING SMALL GROUP COLLABORATION

Dillenbourg and Schneider [54] define collaborative learning as involving situations in which two or more participants interactively build a joint solution to a problem, and distinguish collaborative activity from activities in which tasks are divided and solved independently by individual group members. Similarly, Johnson and Johnson [55] define collaborative groups as follows:

> "A small group may be defined as two or more individuals who (a) interact with each other, (b) are interdependent, (c) define themselves and are defined by others as belonging to the group, (d) share norms concerning matters of common interest and participate in a system of interlocking roles, (e) influence each other, (f) find the group rewarding, and (g) pursue common goals."

In collaborative learning, the common goals are educational and generally culminate in the creation of an educational product. Small group collaborative learning has been shown to result in higher achievement, less stress and greater student satisfaction, and greater appreciation for diversity [20, 21, 55, 56, 57, 58]. Some educators suggest that it may be particularly important and well suited to the online environment as a way of incorporating the social aspects of learning into a virtual environment [28]. Indeed, there is research which suggests that collaborative learning may be very effective online [9, 10, 17, 59]. For example, Hoag and Baldwin [60] found that students learned more in an online collaborative class than in a face-to-face classroom comparison, but that they also acquired greater experience in teamwork, communication, time management, and technology use. On the other hand, some research also suggests collaborative online learning must be carefully managed to be successful when small group projects are employed [45, 61].

Careful structuring of small group collaborative work is particularly important online for several reasons, the most obvious being it is much easier for group members to avoid participating, they can more easily "disappear" online, and much harder to negotiate collaborative activities in the asynchronous online

environment because the give and take of negotiation is extended in time. Online students can just not participate in group activities by showing up for class. Structuring of small group activities must begin, as discussed in relation to collaborative discussion, with a careful identification of the learning goals one wishes to achieve through small group collaboration. Most educators agree that these must include both content learning and collaborative skills, process goals as well as specific products, and goals for both individuals and groups. Each goal should be clearly specified and assessment procedures developed to measure each one. Ideally, such assessment procedures can be embedded in the small group activities themselves and intrinsic to them. Issues surrounding the development of these are discussed below.

First, learning the course content must be an outcome of small group work or why include it, and so it is important to carefully consider what kinds of content can best be learned collaboratively. Students learning to solve problems in a range of content areas can benefit from considering multiple approaches to solutions by working toward collaborative ones. Students learning research and writing techniques can similarly benefit from collaborative endeavors. Such activities can be assessed by assessing their products, but it often helps to break larger tasks into smaller pieces that are also assessed.

At the same time, learning to collaborate with others is an important skill in itself. Thus it is important to consider what collaborative skills one considers most important and develop ways of assessing them. Curtis and Lawson [58], for example, identified the following behaviors as supportive of collaboration:

- giving and receiving help and assistance
- exchanging resources and information
- explaining or elaborating information
- sharing knowledge with others
- giving and receiving feedback
- challenging others contributions
- advocating increased effort and perseverance among peers
- monitoring each others' efforts and contributions.

These behaviors can be noted and assessed by instructors monitoring the discussion.

Other educators consider functions within groups such as discussion leader, facilitator, reporter, observer, and participant as critical and so assign and rotate these roles among individual students and develop separate assessments for each role [62]. It may even be possible to identify specific collaborative skills that are especially important in particular domains [62]. In any case, it is important to value collaborative skills to encourage students to learn to use them.

In addition, most educators agree that it is critical, especially for students collaborating online, that a group identity, a sense of community, be established before serious collaborative group work commences [36, 63]. Thus, many recommend ice-breaker activities that are fun and encourage self-revelation, and which are assessed solely for participation [51]. Another way to help student groups establish community is to explicitly initiate them to the processes and etiquette of online collaboration. Many online courses and programs have required orientation modules that do this.

Second, it is important to assess both the processes and the products of collaboration [64]. For example, when Southern Cross University decided to offer an undergraduate course on group processes online, the

course developers wanted to retain several group activities with specific learning outcomes from the original course. They also required individual group members to complete an observation report on interactions among group participants, a reflective journal that recorded their own perceptions of the experiences, and an essay evaluating the effectiveness of their online group, all of which were assessed. Researchers found that the online experience provided an opportunity for students to learn collaboratively that was equal to that provided in the face-to-face version of the group processes course, but that the added opportunity to analyze and reflect on their own collaborative processes enhanced students' learning experience [61].

One way to collect both processes and products is through portfolio assessment. Portfolios are student-prepared collections of documents that evidence understanding of important concepts or mastery of key skills by requiring students to organize, synthesize, and communicate their achievements throughout the semester. Several different types of portfolios can be used, but most are variations of students' personalized collections of their work over the entire duration of the course. In the case of the assessment of collaborative group work, students might be asked to provide evidence of their contributions to group projects or reflections on the group process, as well as evidence of learning. Portfolio assessments provide each student with the opportunity to demonstrate their understanding of course material as well as their participation in collaborative processes, and, when used longitudinally, how their understandings change over time in response to others contributions.

Third, to help ensure both group interdependence and individual accountability [20, 21, 57] and so to support collaboration, both group and individual assessments are essential. Freeman and McKenzie [65] maintain that it is clearly not adequate to evaluate a collaborative, project-based assignment by merely rating the quality of the final product of the group; rather, "If our courses have the objective of developing students' capacity to work as part of a team, then we need some means of assessing teamwork in a fair and meaningful way." They recommend using Likert-type scales with explicit rubrics to rate at least two dimensions of the group experience: the functioning of the group as a whole, and the performance of each individual member. Often these ratings are completed by group members with oversight from the instructor. Another option is to make the group assessment the sum of individual participants' assessments on some measure of content learning [66]. This sort of assessment makes group participants responsible for the learning of all members of their group, as suggested by Johnson & Johnson.

In a similar vein, Pelz [51] requires all students to participate in a peer learning section of every module of his courses, either giving or seeking assistance. Pelz also assigns a collaborative Internet research paper with requirements and an assessment scheme that encourage both interdependence and individual responsibility in an interesting way. First, students must propose a topic to the class including supporting Internet resources. Students are then required to assist each other by suggesting additional ideas and Internet resources in a collaborative discussion graded as detailed above. Students then develop a research paper which must meet a series of content and skills oriented criteria and include ideas and resources suggested by their peers. If they do not include both ideas and Internet resources suggested by their peers and cite their sources, students are docked points on their research papers. Students are also given points in this activity for ideas and resources used by others. Finally, student papers are posted and each discussed in a graded collaborative forum. The entire process takes an entire semester and involves all students in each others' research and so in each others' learning.

Similarly, Nachmias, et al. [19] report several kinds of peer evaluation and review in the online graduate courses they studied including reference and response to peers' contributions, formative evaluation of peers' work in progress and selection tasks involving judging peers' contributions, with corresponding assessment schemes. They also note a knowledge construction activity, the collaborative creation of an

annotated database of educational websites, in which students were required to find websites meeting specific criteria and then evaluate each others' contributions using a common instrument based on those criteria. Students were assessed on each of these activities as well as on the relevance and completeness of the data base as a whole.

Still another study by Naidu et al. [14] reports significant learning among students participating in a role-play simulation on world politics in which final grades were based on the writing of individual role profiles (25%), the degree of their participation (25%), the quality of their participation (25%) and on a final summary of their simulation activity (25%). Though quite different forms of assessment were employed in these different examples of small group collaborative learning, in all of them both individual and collaborative efforts were assessed to ensure successful collaboration.

IV. COLLABORATIVE EXAMINATIONS

Finally, there are some who argue that the constructivist nature of collaborative learning suggests that the experience might be enhanced by collaboratively designed assessments. For example, Fisher, et al. [62] report that adding collaboratively developed small group activities enforced by the assessments of group members to an online course increased student satisfaction with and learning from collaborative small group projects. Nachmias, et al. [19] describe their successful use of what they call "emergent collaboration" in which small student groups collaboratively develop both the specific activities they will use to solve a particular problem and the ways they will be assessed. Similarly, Murphy, Mahoney and Harvell [68] collaboratively developed group contracts for assessing project-based group work in an online course on telecommunications in education helped build community among groups.

During the last several years, a series of field studies at NJIT have explored the participation of students in all phases of online exam processes. In the initial studies, an assessment procedure termed the "participatory" exam process was developed and piloted (termed "collaborative exams" in publications on this work before 2005). In the participatory examination, individual students participate in creating essay questions, answering questions in a public online discussion space, and suggesting grading for the answers of other students to the questions they created. The results have been generally positive; the majority of students reported learning from creating questions and grading answers, as well as from studying for and answering questions [68, 69, 70, 71, 72]. However, a substantial proportion of students were concerned about the fairness of individual students grading other students, even though the instructor gave detailed rubrics for grading and reviewed the grading.

Extending previous studies, Shen [25] further applied constructivism and collaborative learning theories to assessment, and designed and studied the collaborative online exam. A collaborative exam is an online exam in which small groups of 3–5 students create questions; other individual students answer these questions; the small group grades the answers to the questions they created, using a set of detailed rubrics for grading; and then the instructor reviews the suggested grading and rationale and assigns the final grade. At all phases, the various questions and answers and grading critiques are visible to all members of the class, with assigned pen names used to hide the identity of the question and answer authors. The *participatory online exam* process features similar procedures except that students' involvement in every phase of the exam is individual. Contrasts between the traditional teacher-centered exam, the participatory exam, and the collaborative examination processes are illustrated in Figure 4 below.

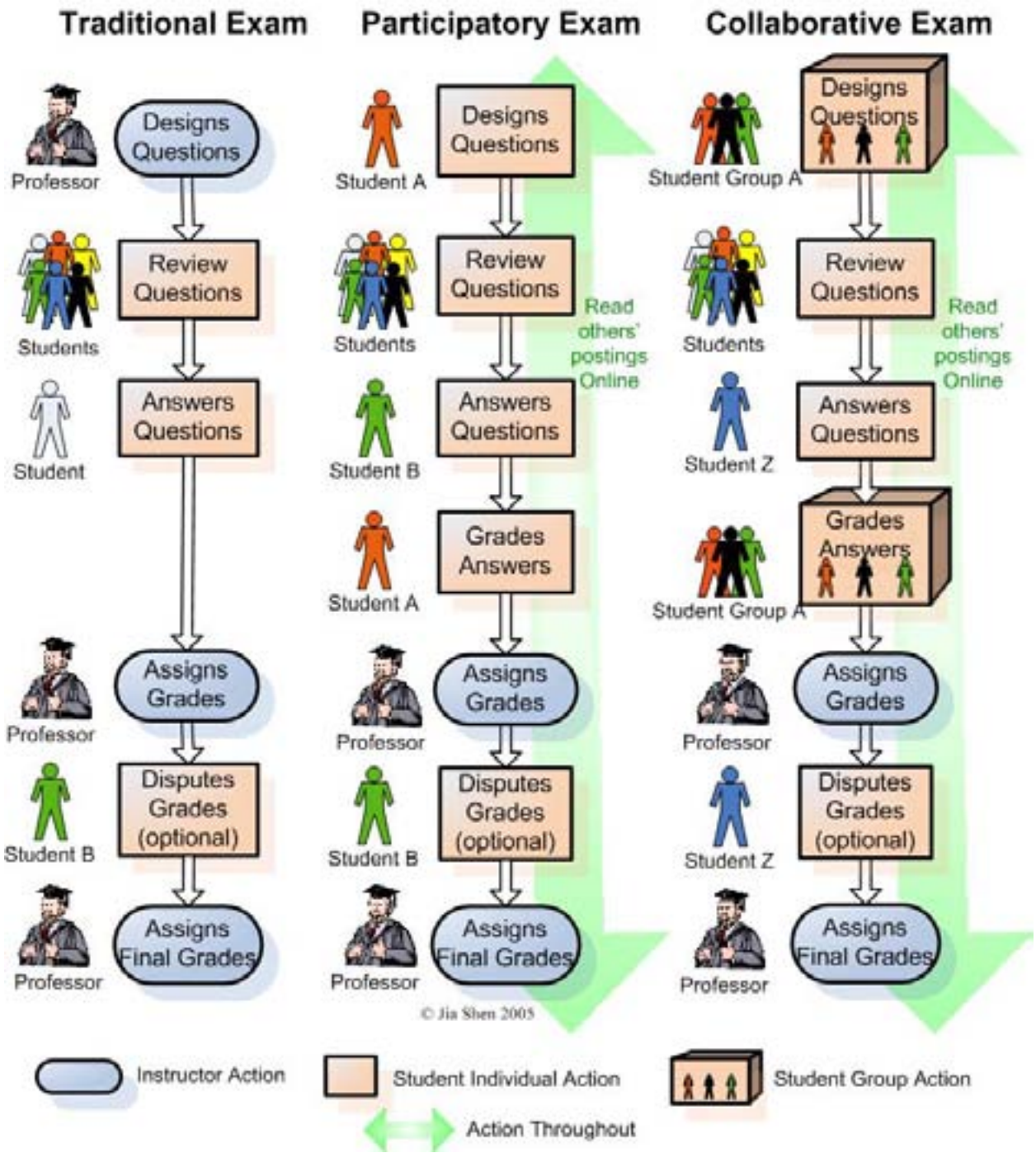

Figure 4: Comparison of Exam Modes (Shen, 2005)

As shown in Figure 4, the collaborative and the participatory exams share the same sequence of steps. The differences between the two online exam modes are in the question design and grading steps. While students in the participatory exam participate in these steps individually, students in the collaborative exam first participate individually (i.e., individually design questions and grades), and then the group as a whole discusses, improves, and reaches consensus on the questions and grades.

During pilot studies, it was determined that extremely detailed instructions needed to be given to students to be able to follow a step-by-step process for this unfamiliar new procedure, the collaborative examination. Appendix A shows the final set of grading rubrics instructions that was developed and used for the grading phase of the process, and which were made available to the students before they composed their questions or answers. As shown in Appendix A, in addition to the answer grade, the question quality and the grading quality in the online exams were also evaluated and graded. Studies on collaborative learning and collaborative assessment show that students are most motivated when each and every part of their effort is recognized [73, 74]. To motivate students' participation in all the steps in the online exam modes, students were graded not only on their answers, but the quality of their questions, and the grades they provide. In most cases the percentage for the three grades in the online exam modes was 15% for question design (question grade), 70% for answer (answer grade), and 15% for quality of grade (grading grade). The grading rubrics in Appendix 1 can serve as a model for other instructors who may wish to adopt the collaborative examination.

A field experiment was then conducted which compared the traditional examination process with the participatory and collaborative examination processes. Results of the 1*3 (exam mode as the independent variable with three conditions) field experiment [25] show that collaborative examinations significantly enhanced interaction and promoted higher order learning. In particular, small group activities in the online learning process significantly increased interactions among students which enhanced their sense of an online learning community. Active involvement in the online exams significantly reduced the use of surface learning in exam study. Overall, students reported significantly higher perceptions of learning in the collaborative exam than the other exam modes.

V. CONCLUSIONS

In this paper, we have argued that the assessment of collaborative learning in online courses is critical to its success. We have acknowledged that assessing collaborative learning is difficult because it requires radically rethinking traditional evaluation techniques. However, we believe such rethinking is also critical because collaboration among students has been repeatedly shown to enhance the effectiveness of online learning. In the body of the paper, we have explored issues surrounding the assessment of three categories of online collaborative learning—collaborative discussion, collaboration in small groups, and the collaborative design of assessments themselves—and given suggestions for developing such measures as well as examples of evaluations successfully used in online courses. Across these discussions, a general approach to developing assessments of collaborative learning can be discerned.

The first step in the process involves specifying the explicit learning goals one wants to achieve through collaborative activities. It is important here to choose goals deemed critical to the success of the desired learning activity, and to consider both content learning and the development of collaborative skills, process and outcome goals, individual and group learning. The next step is to identify specific behaviors that can provide evidence of achieving the selected goals and to assign values to each corresponding to their perceived importance to the collaborative activity. One good way to do this is to develop rubrics which provide finely-detailed characterizations of student performance of each identified behavior at multiple levels with differential values assigned to them. If you choose to have students develop their own assessments, it is important to help them create equally explicit evaluation criteria, perhaps by supplying them with templates to support such development. In that vein, it is terribly important to clearly communicate assessment procedures to students through course documents available at the beginning of a course and accessible throughout it, and through ongoing and timely feedback using the criteria outline in these documents.

In sum, online learning changes not only the nature of teaching and learning, but also the nature of effective assessment processes. Learning and student assessment are not two distinct phases of the course process, but rather, assessment not only guides and motivates the learning, but also can be part of collaborative learning and community building in ALN. Recognizing this evolution in assessment practices, documenting and sharing procedures and rubrics that work and conducting empirical research to evaluate the relative effectiveness of different online practices for assessing collaborative learning, should play a prominent role in ALN research.

VI. REFERENCES

1. **Erwin, T. D.** *Assessing Student Learning and Development*, 14–19. San Francisco: Jossey-Bass, 1991.
2. **Angelo, T. A.** *Definition of assessmen*t. AAHE Bulletin, November 7, 1995.
3. **Entwhistle, N.** Promoting deep learning through teaching and assessment. *Assessment to Promote Deep Learning: Insights from AAHF's 2000 and 1999 Assessment Conferences*, 9–20. Washington DC: American Association of Higher Education, 2001.
4. **Bloom, B.** *A Taxonomy of Educational Objectives*. New York: D. McKay Co, 1956.
5. **Deming, W. E.** *Out of the Crisis*. Cambridge, MA: MIT Press, 1986.
6. **Olson, L.** Quality-management movement spurs interest in new awards for education. *Education Week 11*(8): 1992.
7. **Wright, B. D.** More art than science: the postsecondary assessment movement today. In J. Bourne and J.C. Moore (eds.), *Elements of Quality Online Education: Into the Mainstream, Vol. 5 in the Sloan-C Series*, 185–197. Needham, MA: Sloan-C, 2004.
8. **Huba, M. E. and J. E. Freed.** *Learner-Centered Assessment on College Campuses: Shifting the Focus from Teaching to Learning.* Boston: Allyn & Bacon, 1999.
9. **Scardamalia, M. and C. Bereiter.** Computer support for knowledge building communities. *The Journal of the Learning Sciences* 3(3): 265–283, 1994.
10. **Linn, M. and S. Hsi.** *Computers, Teachers, Peers, Science Learning Partners*. Mahwah, NJ: Erlbaum, 1998.
11. **Pea, R.** Seeing what we build together: distributed multimedia learning environments for transformative communications. *The Journal of the Learning Sciences* 3(3): 285–299, 1994.
12. **Isenhour, P. L., J. M. Carroll, D. C. Neale, M. B. Rosson, and D. R. Dunlap.** The virtual school: An integrated collaborative environment for the classroom. *Educational Technology & Society* 3(3): 74–86, 2000.
13. **Ben-Haym, A.** MOO virtual textual environment and its implications for collaborative learning. Unpublished MA Thesis, Tel-Aviv University School of Education, 1999.
14. **Naidu, S. A. Ip, and R. Linser.** Dynamic goal-based role-play simulation on the Web: A case study**.** *Educational Technology & Society* 3(3): 190–202, 2000.
15. **Lamaster, K. and D. Tannehill.** Preservice teachers as mentors using telecommunications. *International Journal of Educational Telecommunications* 5(1): 25–46, 1999.
16. **Benbunan-Fich, R. & Hiltz, S. R.** Impact of asynchronous learning networks on individual and group problem solving: A field experiment. *Group Decision and Negotiation* 8: 409–426, 1999.
17. **Hiltz, S. R.** Impacts of college-level courses via asynchronous learning networks: Some preliminary results. *Journal of Asynchronous Learning Networks* 1(2): 1997.
18. **Hammond, M.** Issues associated with participation in online forums: The case of the communicative learner. *Education and Information Technologies* 4(4): 354–367, 1999.

19. **Nachmias, R., D. Mioduser, A. Oren, and J. Ram.** Web-supported emergent-collaboration in higher education courses. *Educational Technology & Society* 3(3): 94–104, 2000.
20. **Johnson, D. W. and R. Johnson.** Computer-assisted cooperative learning. *Educational Technology* 26(1): 12–18, 1989.
21. **Johnson, D. W. and R. Johnson.** Positive interdependence: key to effective cooperation. In R. Hertz-Lazarowitz and N. Miller (Eds), *Interaction in Cooperative Groups: The Theoretical Anatomy of Group Learning*, 174–199. Cambridge: Cambridge University Press, 1992.
22. **Kwok, R. C. W. and J. Ma.** Use of a group support system for collaborative assessment. *Computers & Education* 32(2): 109–125, 1999.
23. **Wilson, E. V.** Examnet Asynchronous Learning Network: augmenting face-to-face course with student-developed exam questions. *Computers & Education* 42: 87–107, 2004.
24. **Shindler, J.** Examining the soundness of collaborative essay exams in teacher education courses. *National Forum of Teacher Education Journal* 12(3): 2003.
25. **Shen, J.** *Collaborative Examinations in Asynchronous Learning Networks: Field Experiments On Collaborative Learning Through Online Assessments*. Ph.D. dissertation, Department of Information Systems, NJIT, May, 2005. Online: http://www.alnresearch.org.
26. **Topping, K. J.** Peer assessment between students in college and university. *Review of Educational Research* 68(3): 249–276, 1998.
27. **Sluijsmans, D. M. A., S. Brand-Gruwel, J. Van Merriënboer, and T. R. Bastiaens**. The Training of Peer Assessment Skills to Promote the Development of Reflection Skills in Teacher Education. *Studies in Educational Evaluation* 29(1): 23–42, 2003.
28. **Harasim, L.** *On-line Education: Perspectives on a New Environment.* New York: Praeger, 1990.
29. **Levin, J. A., H. Kim, and M. M. Riel.** Analyzing Instructional Interactions on Electronic Message Networks. In L. Harasim (Ed.), *On-line Education: Perspectives on a New Environment.* New York: Praeger, 1990.
30. **Hiltz, S. R.** *The Virtual Classroom: Learning without Limits via Computer Networks.* Norwood, NJ: Ablex, 1994.
31. **Poole, D. M.** Student participation in a discussion-oriented online course: A case study. *Journal of Research on Computing in Education* 33(2): 162–177, 2000.
32. **Garrison, D. R.** Cognitive presence for effective asynchronous online learning: The role of reflective inquiry, self-direction and metacognition. In J. Bourne & J. C. Moore (Eds.), *Elements of Quality Online Education: Practice and Direction*, 47–58. Needham, MA: Sloan-C, 2003.
33. **Walther, J.** Interpersonal effects in computer mediated interaction. *Communication Research* 21(4): 460–487, 1994.
34. **Gunawardena, C.** Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications* 1(2/3): 147–166, 1995.
35. **Rourke, L., T. Anderson, D. R. Garrison, and W. Archer.** Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education* 14(2): 2001. Online: http://cade.icaap.org/vol14.2/rourke_et_al.html.
36. **Rovai, A. P.** A preliminary look at structural differences in sense of classroom community between higher education traditional and ALN courses. *Journal of Asynchronous Learning Networks* 6(1): 41–56, 2002.
37. **Swan, K.** Building communities in online courses: The importance of interaction. *Education, Communication and Information* 2(1): 23–49, 2002.

38. **Gunawardena, C. and F. Zittle.** Social presence as a predictor of satisfaction within a computer mediated conferencing environment. *American Journal of Distance Education* 11(3): 8–26, 1997.
39. **Wegerif, R.** The social dimension of asynchronous learning networks. *Journal of Asynchronous Learning Networks*,2(1): 34–49 ,1998.
40. **Tu, C. H.** On-line learning migration: from social learning theory to social presence theory in CMC environment. *Journal of Network and Computer Applications* 23(1): 27–37, 2000.
41. **Picciano, A. G.** Beyond student perceptions: Issues of interaction, presence and performance in an online course. *Journal of Asynchronous Learning Networks* 6(1): 21–40, 2002.
42. **Richardson, J. C. and K. Swan.** Examining social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks* 7(1): 68–88, 2003.
43. **Hawisher, G. E. and M. A. Pemberton.** Writing across the curriculum encounters asynchronous learning networks or WAC meets up with ALN. *Journal of Asynchronous Learning Networks* 1(1): 1997.
44. **Jiang, M. and E. Ting.** A study of factors influencing students' perceived learning in a web-based course environment. *International Journal of Educational Telecommunications* 6(4): 317–338, 2000.
45. **Swan, K., P. Shea, E. Fredericksen, A Pickett, W. Pelz, and G. Maher, G.** Building knowledge building communities: Consistency, contact and communication in the virtual classroom. *Journal of Educational Computing Research* 23(4): 389–413, 2000.
46. **Swan, K.** Virtual interactivity: Design factors affecting student satisfaction and perceived learning in asynchronous online courses. *Distance Education* 22(2): 306–331, 2001.
47. **Jaffee, D.** Asynchronous learning: Technology and pedagogical strategy in a computer-mediated distance learning course. *Teaching Sociology* 25(4): 262–277, 1997.
48. **Vinaja, R., and M. Raisinghani.** Analysis of strategies used in teaching an online course in a predominantly hispanic university. *JCSC*, March: 70–79, 2001.
49. **Saltz, J., S. R. Hiltz, and M. Turoff.** Student Social Graphs: Visualizing a Student's Online Social Network. *Proceedings, CSCW '04*, 2004.
50. **Zumbach, J. and P. Reimann.** Combining Computer Supported Collaborative Argumentation and Problem-Based Learning: an Approach for Designing Online Learning Environments, 1999. Online: http://d3e.open.ac.uk/cscl99/Zumbach/Zumbach-paper.html.
51. **Pelz, W.** (My) Three principles of effective online pedagogy. *Journal of Asynchronous Learning Networks* 8(3): 33–46, 2004.
52. **Ho, C. H.** Assessing Electronic Discourse: A Case Study in Developing Evaluation Rubrics. Paper presented at the 14th Annual Meeting of the Society for Text and Discourse (ST&D). Chicago, August 1–4, 2004.
53. **Grice, H. P.** *Studies in the Way of Words*. Cambridge, MA: Harvard University Press, 1989.
54. **Dillenbourg, P. and D. Schneider.** Collaborative Learning and the Internet. ICCAI 95, 1995. Online: http://tecfa.unige.ch/tecfa/research/CMC/colla/iccai95_1.html.
55. **Johnson, D. W. and R. Johnson.** *Joining Together: Group Theory and Group Skills, Fifth Edition.* Boston: Allyn & Bacon, 1994.
56. **Pressley, M. and C. R. McCormick,** *Advanced Educational Psychology for Educators, Researchers, and Policy Makers.* New York: Harper Collins, 1995.
57. **Slavin, R. E.** *Cooperative Learning*. Boston: Allyn & Bacon, 1995.
58. **Curtis, D. D. and M. J. Lawson.** Exploring collaborative online learning. *Journal of Asynchronous Learning Networks* 5(1): 21–34, 2001.

59. **Gunawardena, C. N., Lowe, C. A. & Anderson, T.** Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research* 17(4): 397–431, 1997.
60. **Hoag, A. and T. F. Baldwin.** Using case method and experts in inter-university electronic learning teams. *Educational Technology and Society* 3*(*3), 337–348, 2000.
61. **Sturgill, A., W. Martin, and G. Gay.** Surviving technology: A study of student use of computer-mediated communication to support technology education. *International Journal of Educational Telecommunications* 5(3): 239–259, 1999.
62. **Fisher, K., Phelps, R. and A. Ellis**. Group processes online: Teaching collaboration through collaborative processes. *Educational Technology & Society* 3(3): 484–495, 2000.
63. **Palloff, R. and K. Pratt.** *Building Learning Communities in Cyberspace: Effective Strategies for the Online Classroom.* San Francisco: Jossey-Bass, 1999.
64. **Macdonald, J.** Assessing online collaborative learning: Process and product. *Computers and Education* 40(4): 377–391, 2003.
65. **Freeman, M. and J. McKenzie.** SPARK, a Confidential Web-based Template for Self and Peer Assessment of Student Teamwork: Benefits of Evaluating Across Different Subjects. *British Journal of Educational Technology* 33(5): 551–569, 2002.
66. **Boud, D., R. Cohen, and J. Sampson.** Peer learning and assessment. *Assessment and Evaluation in Higher Education* 24(4): 413–426, 1999.
67. **Murphy, K. L., S. E. Mahoney, and T. J. Harvell.** Role of contracts in enhancing community building in web courses. *Educational Technology and Society* 3(3): 409–421, 2000.
68. **Shen, J., K.-E. Cheng, Y. Cho, S. R. Hiltz, and M. Bieber.** Evaluation of an on-line collaborative examination process. *Proceedings of the 2000 Americas Conference on Information Systems*, CD ROM, 2000.
69. **Shen, J., S. R. Hiltz, K.-E. Cheng, Y. Cho, Y. and M. Bieber.** Collaborative examinations for asynchronous learning networks: Evaluation results. *Proceedings of the 34th Hawaii International Conference on Systems Sciences*, Maui, Hawaii, IEEE Computer Society Press, 2001.
70. **Shen, J., K.-E.Cheng, ,M. Bieber, and S. R. Hiltz.** Traditional in-class Examination vs. collaborative online examination in asynchronous learning networks: Field evaluation results. *Proceedings, Americas Conference on Information Systems 2004*, New York City, NY.
71. **Shen, J., S. R. Hiltz, and M. Bieber.** Collaborative assessment in asynchronous learning networks: Research in progress. In J. Bourne and J. C. Moore (eds.). *Elements of Quality Online Education: Engaging Communities, Volume 6* in the Sloan-C Series. Needham, MA: Sloan-C, 2005.
72. **Wu, D., M. Bieber, S. R. Hiltz, S. R. and H. Han**. Constructivist learning with participatory examinations. Proceedings of the HICSS 37th Hawaii International Conference on Systems Sciences, Big Island, Hawaii, IEEE Computer Society Press, CD-ROM, 2004.
73. **Bruffee, K. A**. Collaborative learning: higher education. *Interdependence and the Authority of Knowledge*. Baltimore, MD: John Hopkins University Press, 1999.
74. **Corbitt, G., L. Wright, L. and B. Martz.** Addressing the challenges of the future: Implementing a collaborative student environment at a university business school. *Proceedings of the 32nd Hawaii International Conference on System Science*, 1999.

VII. ACKNOWLEDGMENTS

NJIT's research relating to assessment in online courses has been partially supported by the Alfred P. Sloan Foundation and the New Jersey Center for Pervasive Information Technology (NJPIT); and the National Science Foundation (0534520 and 0454081); the opinions presented here are those of the authors

and not necessarily of the sponsors. Portions of this paper are adapted from the Ph.D. thesis by Jia Shen (2005) and from a preliminary report of research in progress (Shen et al, 2004).

VIII. ABOUT THE AUTHORS

Starr Roxanne Hiltz is Distinguished Professor, College of Computing Sciences, New Jersey Institute of Technology. She conducts research on the applications and social impacts of computer technology. Her research interests currently include Group Support Systems, Asynchronous Learning Networks, and Pervasive Computing. One current project is the *WebCenter for Learning Networks Effectiveness Research*, a collaborative knowledge base about research on ALN (see http://www.ALNresearch.org). Her most recent book, co-edited with Ricki Goldman, is *Learning Together Online: Research on Asynchronous Learning Networks* (Erlbaum, 2005). For more information, see http://eies.njit.edu/~hiltz.

Jia Shen is an Assistant Professor in the School of Management at New York Institute of Technology. She completed her Ph.D. in the Information Systems Department at New Jersey Institute of Technology in May 2005. Her current research interests include Computer Supported Cooperative Work (CSCW), Asynchronous Learning Networks (ALN), Human Computer Interaction, and Knowledge Management. She has presented papers at many conferences including AMCIS, HICSS, and CHI. She may be reached through http://iris.nyit.edu/~jshen/.

Karen Swan is Research Professor in the Research Center for Educational Technology at Kent State University. Dr. Swan's research has been focused mainly in the general area of media and learning on which she has published and presented nationally and internationally. Her current research focuses on online learning, and on student learning in ubiquitous computing environments. Dr. Swan has also authored several hypermedia programs, co-edited a book on *Social Learning from Broadcast Television* and is currently working on a co-edited book and a DVD ROM on ubiquitous computing. She is an Effective Practices Editor for the Sloan Consortium, the Special Issues Editor for the *Journal of Educational Computing Research*, and Editor of the *Journal of the Research Center for Educational Technology (*http://www.rcetj.org).

IX. APPENDIX A: GRADING RUBRICS FOR THE COLLABORATIVE EXAM

1. Question Grading Criteria—used by the professor

(Total 15 points—group grade, except specially advised by the instructor)

Quality	• 9 points: Questions cover different aspects of course material (3 points), are within the exam scope (3 points), and are identified with difficulty levels (3 points). • 4 points: Clarity of questions, including quality of the writing.
Following Directions	• 2 points: Submitting in the correct place and format in the WebCT conference. (Postings not submitted anonymously, or not submitted in the correct place, will be deducted 2 points) • Late submission: minus 4 points for submissions within 24 hrs past the deadline; minus 8 points for submissions past 24 hours within 48 hrs of the deadline; no submissions will be accepted after 48 hrs.

2. Answer Grading Criteria—used by students and the professor

(Total 70 points (35 points each answer)— Individual grade)

Quality	• 23 points: The correctness and completeness of the answer, including citing most of all relevant course materials, considering all sides of issues, synthesizing etc. • 7 points: The quality and clarity of writing, including providing justification to points, etc.
Following Directions	• 3 points: Following editing guidelines including correct citation format (deduct up to 3 points), and length (deduct 3 points if the answer is under 750 words or exceeds the 1,700 words limit.) • 2 points: Submission of the answer in the correct place and format • Late submission: minus 10 points for answers submitted within 24 hrs past the deadline; minus 20 points for answers submitted after 24 hrs within 48 hours of the deadline; no answer will be accepted after 48 hrs.

Plagiarism: All answers must be examined for plagiarism by considering: 1) whether the answer gives proper citation to the source of information; 2) whether the writing shows the student's own understanding of the knowledge. Proper citation must be used every time sentences or paragraphs are copied from books, papers, or other resources. Students should demonstrate their own understanding of the knowledge by explaining in their own words concepts, theories, methods, and/or providing summaries, examples, etc. If you think parts of the answer are plagiarized, please post the supporting material in your answer grading and grade appropriately.

3. Grade justification Grading Criteria—used by the professor

(Total 15 points—group grade, except specially advised by the instructor)

Quality	• 9 points: Quality of grading, including providing a full written explanation (justification) of the grading with at least 3 full sentences explanation for each of the grading categories. • 4 points: Clarity of justifications.
Following Directions	• Same as the question grading criteria