Assessment & Evaluation in Higher Education
2009, 1–22, iFirst Article

Routledge
Taylor & Francis Group

Chemistry Is in the news: assessing intra-group peer review

Kathleen M. Carson[a]* and Rainer E. Glaser[b]*

[a]*Department of Educational Leadership and Policy Analysis, University of Missouri, Columbia, Missouri 65211, USA;* [b]*Department of Chemistry, University of Missouri, Columbia, Missouri 65211, USA*

Interdisciplinarity is rapidly becoming a norm within both the professional and academic worlds, and the ability to collaborate is becoming an essential skill for all graduates. *Chemistry Is in the News* (*CIITN*) is a curriculum that aims to teach students this skill by engaging student collaborative groups in a project that ties real world events and topics to the content taught in the classroom. While the collaborative activity has been successful in many ways, the challenge of maintaining individual accountability within the collaborative activity has persisted. The need to balance the tension between promoting collaboration and maintaining individual performance standards drove the development of an intra-group peer review system. In developing this peer review system, four goals guide the design: the desire to promote collaboration, to produce a differentiated score among group members reflecting the contribution each person made, to improve student perception of fairness and accuracy in the assessment process of *CIITN* and to avoid artificially inflating students' grades. The system was assessed in the winter semester of 2004 in a large lecture course at a major Midwestern university via student questionnaires and the *CIITN* scores. Evidence is provided to suggest that the intra-group peer review system has met its core goals.

Keywords: science education; interdisciplinarity; group collaboration; peer review; accountability

The search for more effective teaching methods particularly in large lecture classes has led to the wide adoption of collaborative group work (Cooper 2005; McKeachie 2002; Smist 2004; Woolfolk 2004). Collaborative group work has been touted as providing a support network for students and increasing student retention (Gupta 2004; Tinto 1997), providing additional cognitive benefits (Bransford, Brown, and Cocking 2000) and increased learning (Slavin 1996). It is often integrated as a major component of constructivist learning strategies, which focus on linking the subject material to the students' own experiences, so they gain a deep understanding of the subject (Colburn 2000; Yager and Lutz 1994).

While collaborative group work undoubtedly provides many benefits for students and an opportunity for instruction to engage in new and creative methods, it does present a number of challenges including the development of meaningful activities that are appropriate for group work, resulting in significantly different products as a

*Corresponding author. Email: kathleen.carson@gmail.com; glaserr@missouri.edu

ISSN 0260-2938 print/ISSN 1469-297X online

DOI: 10.1080/02602930902862826
http://www.informaworld.com

consequence of the students' collaboration (Lotan 2003), the organisation of students into productive groups and the assessment of individual students within the group (Conway and Kember 1993; Goldfinch and Raeside 1990; Kruck and Reif 2001). This last challenge is particularly salient in large courses in which individual assessment of students is necessary but where there is very limited opportunity for the teaching staff to adequately observe or interact with students to determine how well they contribute to groups' final products.

In fact, one of the greatest barriers to implementing collaborative work into science courses has been student resistance to alternative teaching methods (Carlone 2004; Searby and Ewers 1997; Seymour 2002). Large courses at large research universities that are populated by students who are accustomed to a competitive atmosphere and succeeding via traditional methods, precisely the setting into which *CIITN* has been integrated, have the highest levels of opposition according to Seymour (2002, 98). This opposition manifests itself in students' lack of willingness to participate in class, students' negative course evaluations and even in students appealing to the relevant administration for redress. Seymour points out that failure to address the students' dissatisfaction, or more to the point, the roots of this attitude, can result in substantial professional consequences for the faculty members implementing the new curriculum (p. 97). This is despite the evidence that the vast majority of students who leave science, mathematics and engineering do so because they are not meaningfully engaged by the standard pedagogy (Seymour and Hewitt 1997). Thus, the design and development of a curriculum based on collaborative work must take into account both the need to engage students meaningfully while allaying their fears about their performance and in the case of group work, freeloading (Ballantyne, Hughes, and Mylonas 2002).

A variety of solutions to assess students as individuals working in a collaborative setting have been proposed: having students submit parts of the project independently (Goldfinch and Raeside 1990), having groups report on the group function (Boud, Cohen, and Sampson 1999) or having students report individually on the group function. Each of these has its own weaknesses, ranging from undermining the purpose of group work in the case of dividing up project submission to failing to communicate to students the importance of their reflections on group work by not adequately linking the report on function to a final score.

Another common solution is to have students engage in some form of intra-group peer review, asking them to reflect on and then assess the quality of the various members' contributions to the workings of the group and the final project, providing some sort of score that will clearly contribute to a final score. This method is supported by a large body of evidence from the fields of sociology and psychology. Such a system is necessary to prevent or at least reduce freeloading because, as Olson (1982) asserted, without selective incentives, a rational, self-interested individual will not act in the group's interest. However, individuals will punish non-cooperative behaviour of a partner, given the opportunity (Fehr and Gächter 2000, 2002). In addition, such a system encourages critical thinking skills and introspection, contributing to lifelong learning (Strom and Strom 1998). Students also gain experience in evaluating the work of others, an important part of professional development for anyone entering the science community specifically or the professional world more generally, as experience in peer review correlates with improved peer review reliability and validity (Marsh, Jayasinghe, and Bond 2008). This final solution has been employed with the curriculum of *CIITN*.

Chemistry Is in the news in the classroom

Chemistry Is in the News (*CIITN*) is a curriculum that has been in development and implemented since 1997 (Glaser 2003; Glaser and Carson 2005; Glaser and Poole 1999; Hume et al. 2006). In the winter semester of 2004, *CIITN* was employed in the first semester of a two-semester organic chemistry course sequence for science majors at the University of Missouri, Columbia. Both courses of the sequence are offered every winter and fall semesters, however, students who are 'on schedule' would take the first course in the sequence in the fall. This course has an enrolment with approximately 230 students.

The full *CIITN* curriculum consists of three assignments completed over the course of the semester by collaborative groups of three to five students (Carson et al. 2009). The first assignment is a 200-word abstract of a science-related newspaper article that is assessed by a teaching assistant (TA) through a constructive review, followed by an opportunity for students to revise the abstract, and a final review. The second assignment is the written description of a chemical reaction along with the structures of that reaction drawn using structure drawing software, ChemDraw®, and molecular modelling software, Chem3D®, to show the shape and internal orientation of relevant molecules in the reaction. The second assignment is graded in the same manner as the first. These two assignments serve two purposes. First, the assignments require students to begin working together early in the semester despite having limited content knowledge. Second, students practise the relevant skills such as writing about science for both a broad audience and a specialised audience, and using such software chemists use to communicate their work and thus prepare students for the third and final assignment, the *CIITN* portfolio.

A *CIITN* portfolio consists of a current news article, interpretive comments containing hyperlinks, and questions and answers. Students link the topic of the article to the content they have learned in class in the interpretive comments, providing additional background and information by inserting links to high quality and credible information sources on the Internet. They then further explore the chemistry and social, political and philosophical elements of the topic through questions and answers. The students create portfolios via the *CIITN* webtool (Glaser, Wu, and Sui 2004; Wu and Glaser 2004) which provides a template for students to construct portfolios on the Internet. The student groups then peer review the portfolios using the webtool, and also going through a constructive review round, revision and final review (Carson, Hodgen, and Glaser 2006; Glaser and Carson 2005). Following the completion of the portfolio, students complete intra-group peer review using the webtool.

During the winter semester 2004, the three *CIITN* projects accounted for 200 points of a total of 850 points or 23.5% of the overall course grade. The first and second assignments were worth 50 points each and the scores were entirely based on the score the TA assigned to the work. The third assignment, the *CIITN* portfolio, was worth 100 points and combined the portfolio score and the intra-group peer review score.

Intra-group peer review design

The instructor incorporated intra-group peer review into the *CIITN* curriculum based on student feedback about the integration of *CIITN* into the course. Initially, student groups were required to submit Group Dynamics Reports to the instructor via email.

These reports had no impact on the students' project or overall course grade except in cases when the group reported that a group member made no contribution to the project. In the departmental teaching evaluations, students expressed dissatisfaction with all students receiving the same score regardless of the quality or quantity of their contribution to the group's project, a common complaint of students with this scoring method (Barfield 2003; Boud, Cohen, and Sampson 1999). In response to this, the instructor implemented an initial intra-group peer review system, in which each student was asked to assess each group member, assigning each a score out of 50 points. These scores were then averaged to arrive at an independent intra-group peer review score. This system proved to be unsatisfactory because, despite the students' assertion that they would prefer a system that allowed for differentiation, the vast majority of students assigned full points to all group members. It was also problematic that the intra-group peer review score had little validity or reliability across the course. Under this system, members of a group that had met very little and collaborated very minimally could have the same or even better score than members of a group that had met frequently and collaborated to the fullest extent. The consequence of this system is that the intra-group peer review points were, in effect, 50 free points without any substantial link to the work done.

The goal, therefore, was to develop a new intra-group peer review system to achieve more valid individual grades as well as improve student perception of fairness of the grading scheme. Boud, Cohen, and Sampson (1999) offered a number of categories for consideration when developing a grading system for peer learning to achieve these goals by 'establishing congruence between assessment practices and the kinds of learning a course aims to promote' (414). They include a focussing on key outcomes, designing assessment holistically, contributing to lifelong learning, balancing the individual and the group, and balancing the process and the product.

Focus on key outcomes

In order for the system to be effective and meaningful, 'assessment needs to focus on the central outcomes desired as part of education' (Boud, Cohen, and Sampson 1999, 419). In the case of *CIITN*, the assessment should take into account both the subject-matter learning as well as the teamwork elements of the project. Assessment of the projects as a whole does accomplish this in that the three assignments are graded based on their quality, apart from the quality of group work and the intra-group assessment focuses on the teamwork element. However, it is also important to recognise in a content-centred class such as organic chemistry, a vital contribution as well as an important output of the groups' collaboration is mastery of the subject-matter. Therefore, it is important to integrate that as part of the intra-group peer review in a balanced manner. A balanced approach will neither unduly punish a student with more limited content knowledge who can still contribute greatly to the group and who stands to gain much from the collaborative group learning activity nor excessively reward a student who is very strong in the content area but not necessarily strong as a collaborator.

Holistic design

In addition to focussing on key outcomes, the assessment also needs to be integrated into the entire course assessment schema (Boud, Cohen, and Sampson 1999). For a

collaborative group activity to be effectively integrated into a course, the rest of the course must also be supportive of collaboration. To fulfil this need, the courses into which *CIITN* has been integrated, absolute grading, as opposed to curved grading, is the system used. This emphasises reaching standards, content mastery and learning as opposed to relative performance. Thus, it reduces the incentive for students to compete with one another for a limited number of As and Bs and facilitates student–student interaction by making each student's final grade dependent only on their own performance, not on how they performed relative to their classmates.

Contribution to lifelong learning development

The assessment should focus on outcomes that contribute to the ability to engage in lifelong learning. This assessment would include an emphasis on working with others (Boud, Cohen, and Sampson 1999). In addition, the *CIITN* project is designed to develop other skills related to lifelong learning such as information access, goal definition and meaningful assessment of others' work, which also contribute to the group function.

Individual versus the group

The balance must also be struck between assessing the group and assessing the individuals in the group. It is necessary to explore how the assessment can 'foster group learning whilst not inhibiting individual achievement' (Boud, Cohen, and Sampson 1999, 423). This means developing systems that insulate students against peer pressure and freeloading.

Product versus process

A final question to address is whether the focus should be on the process or on the product of the group work. The situation is not necessarily dichotomous, requiring a focus on one or another. However, it is important to establish what is important in the assessment and develop grading criteria accordingly.

Development of the peer review system

Integrating intra-group assessment as part of project assessment

In examining these categories, a number of goals for the *CIITN* assessment are evident. First, because one of the goals of the *CIITN* group activity is both to encourage collaboration and to acquire and demonstrate content mastery, it is necessary to incorporate both the aspects into the overall assessment of *CIITN* activities. Goldfinch and Raeside (1990) offer a mechanism for the combination of these two elements. In their system, students participate in a two-part process of intra-group peer review, first naming who contributed to the major parts of the project and then scoring how each group member performed on a set of skills. These individual sections of scoring are then weighted and added together to arrive at the peer assessment (PA) score. The PA score is then converted to a PA factor with a formula or a table, which is weighted and multiplied by the group project score to arrive at a final score for the individual student (Goldfinch 1994; Goldfinch and Raeside 1990). This solution offers a number of benefits.

First, it encourages both collaboration and high-quality work by communicating to students that the two are, in fact, interdependent. It is necessary to collaborate to produce a high-quality portfolio. Thus, this system focuses on both key outcomes and a balance between process and product. Second, it mitigates score inflation by combining the group and individual score, achieving the proper balance between the individual and the group through the weighting of the PA factor. The rationale is straightforward: quality of the processes of the group is a major contributing factor in the quality of the group's final product and therefore the scores should be linked. However, this rationale is only appropriate when there is adequate feedback and opportunity for revision, as is the case with the two rounds of review for the *CIITN* portfolio assignments. This is also in contrast to the original system employed with *CIITN* in which every student had the opportunity to earn full participation points, regardless of the quality of their final product.

The system put forth by Goldfinch and Raeside (1990) has several limitations however, some of which Conway and Kember (1993) and Goldfinch (1994) expounded. One of these is that the calculation of the PA factor is complex and not sufficiently transparent for the students (Conway and Kember 1993). It is an important pedagogical need that grade calculation is transparent, but it is essential if students are involved in the determination of grades. In addition, the first part of the PA factor, naming of the students, poses several problems. First, it is inaccurate as Conway and Kember (1993) assert in that the number of times a student is named does not necessarily correlate with how they are scored in the second part. Second, it is a form of forced ranking, which discourages collaboration and instead further encourages students to do as much as possible independently so that they have contributed most to the group project, getting a good PA factor, and so that they have greater control over how the project itself looks, getting a good project score. Third, it encourages students to think of the project as discrete blocks instead of an integrated whole. Thus, while the system put forth by Goldfinch and Raeside (1993) achieves some of the goals of *CIITN* assessment, it is necessary to craft an alternative mode of calculation for the PA factor.

Kruck and Reif (2001) offer part of the solution. In the system put forth, students are given a set number of points, not evenly divisible by the number of group members, to assign to their group members as integers as they see fit. The scores each student receives forms the peer evaluation score. Kruck and Reif justify this system as a system that forces students to score competitively by asserting that it requires students to recognize differences in students' contribution by allocating a 'scarce resource' (42), i.e. points. However, competitive scoring does not promote collaboration, though it could be framed differently for students to mitigate such an effect if not avoid it all together and to allow for the equal distribution of points in well-functioning groups in which it would be inappropriate to punish any one student.

Scoring criteria

Aside from determining how scoring is to be logistically accomplished, it is also necessary to determine the criteria. In the *CIITN* curriculum, there are two main factors that must be taken into consideration when establishing grading criteria for students. First, the *CIITN* portfolio, the final assignment, is a fairly open-ended project, which encourages and rewards creativity, exploration and unique solutions both through the formal peer review of projects and with instructor recognition and

bonus points. Therefore, a goal of the scoring criteria is to not overly constrain how students approach producing the portfolio by defining too explicitly how a group should work. Second, because of the increasing number of students who are working to support themselves while at the university, as well as the numerous other activities students necessarily engage in to prepare for life after graduation, such as community service, internships and shadowing, it is also essential to allow for flexibility in how a group agrees to best carry out its work. Thus, when determining what sort of rubric would be used in establishing group grading criteria, it was decided to use a holistic rubric (Hall 2003; Phillip 2002) with the various aspects of group work listed, such as time contribution, expertise and attitude and levels of performance outlined, but without specific points being assigned to any resulting category (see Appendix). This was done as opposed to creating a more rigid rubric with point values assigned to all categories. The more flexible rubric allows groups to use alternative means of creating the portfolio and accommodate group members with time constraints and still have an applicable rubric with which to score their group members.

CIITN intra-group peer review

The *CIITN* intra-group peer review system is a hybrid of the two systems designed by Goldfinch and Raeside (1993) and Kruck and Reif (2001). Students are given the rubric at the beginning of the semester when they form groups of three to five students and the intra-group scoring is roughly outlined for the students. It is explained to students in class that rubric categories are some of the things they should consider when choosing group members. For example, if it is important to them that all group members attend all group meetings, it would be better to form a group made of students who have similar schedules as opposed to automatically forming a group with a former lab partner who works at a job 20 hours per week.

The instructor assigns the first assignment the fourth week of the semester, making it due the sixth week of the semester. The week it is assigned, TAs hold computer training sessions for 15 to 25 students at various times throughout the week to teach students the skills they will need to complete the assignment. The instructor and TAs strongly encourage group members to attend the computer training session together; however, the teaching staff stress it is important that they attend a session even if the students cannot come with their group. The instructor assigns the second assignment the seventh week of the semester with another training session given that week. It is due the eight week of the semester. Two weeks before the first version of the portfolio is due, the TAs hold a final computer training session. Here they explain the process of using the *CIITN* webtool for portfolio creation, portfolio review and intra-group peer review and the philosophy of intra-group peer review.

As opposed to framing scoring as the distribution of a scarce resource as advocated by Kruck and Reif (2001), the teaching staff encourages students to see scoring as communicating how well a student fulfiled the group's expectations. If the student completely fulfils expectations, then they should receive a full share of the 100 points, representing the total group function. If a student went above and beyond the group's expectations, he/she should receive more than a full share, which is possible in this system. However, for one student to have done more than their part, another student would have necessarily not contributed as they should have and should receive less than a full share, making points available to reward the student who contributed more than expected. Should the student feel that one or more students did not contribute

adequately but another student, aside from him or herself, did not compensate for those students, he/she has the option of not assigning all 100 points.

Following the final round of peer review of the portfolio, students login to the *CIITN* webtool (Glaser et al. 2004; Wu and Glaser 2004) as an individual and carry out intra-group peer review. Once logged into their individual accounts, the students can access an electronic copy of the rubric. They then access the intra-group peer review form where all group members except for the individual grading are listed with a point value field and a comment field for each. At the bottom there is a 'Points Left' field. Students must distribute all 100 points between the various fields and provide a justification for that value in the comment field for each student. If fewer than 100 points or more than 100 points are distributed between the group members and the 'Points Left' field, the students receive an error message and must return to the scoring form and modify their entries. In addition, students are required to provide a comment for each group member; an error message is also returned if they fail to fill in one or more comment fields. Requiring comments encourages students to be thoughtful in their scoring, to provide a reference in the case of grade disputes and, most importantly, to give students more meaningful feedback with regard to their collaborative group work skills (Boud, Cohen, and Sampson 1999).

The webtool does contain three safety mechanisms for the intra-group peer review process. First, the students review their group members anonymously. A second mechanism is that students have a score of zero until they have completed their intra-group peer review. This serves two purposes. First, as Boud, Cohen, and Sampson (1999) pointed out, it communicates to students the importance of carrying out the peer review without allocating any additional points to the process, thereby not diluting the entire assessment scheme. Second, it reduces the opportunity and motivation for retaliation. Students cannot see what they have received from their group members and thus they do not have the motivation of payback unless students have otherwise communicated how they are going to grade members. This mechanism also makes it easier for the teaching staff to determine who has completed peer review and who has not yet finished because on the webtool-generated grade sheet, the students who have not completed intra-group peer review have earned a zero on intra-group peer review and thus on the project overall, though a webtool administrator can retrieve a student's scores if necessary.

The third mechanism is to protect students in case one of the group members does not complete intra-group peer review, thereby reducing the total points allocated to each student. If this occurs, the rest of the group members' intra-group peer review scores are not figured as points out of 100, as is the case if all students in the group complete peer review. Instead, the group members' scores are calculated by dividing the points that each student earned by the 'optimal' score possible with one or more group members not participating in peer review. In a group with five students, for example, each student should receive four scores; the optimal score received from each group member would be 25 points and the total optimal score would be 100. In the case of one student not completing peer review, the four students who would subsequently lack an individual intra-group peer review score would now have an optimal total score of 75 points; so those students receive a score of points out of 75. The webtool accomplishes this automatically and the students' scores are displayed using this calculation throughout the peer review period. This has the additional benefit of minimising the number of students who panic when they access their score before everyone has completed intra-group peer

review because their final score usually falls within a normal range even if their raw intra-group score does not.

Assessment

In keeping with the concept of peer review and the motivation behind implementing intra-group peer review, the assessment of the intra-group peer review system is guided by how students use and react to it. With that in mind, with this assessment we aim to answer three key questions: (a) What is the effect of peer review on students' project grade; (b) What are students' perceptions of this system and (c) How supportive of the peer review process do the students perceive the system as being? In answering these questions, we anticipate finding areas in which this tool can be improved either through its structure or its presentation.

Methods

The data for this assessment have been collected via pre- and post-surveys and from the scores recorded from 2004 winter semester. The surveys were distributed during class and a small number of bonus points were awarded for their completion, 10 points for the pre-survey and 20 points for the post-survey. The surveys could be submitted confidentially: students were given the opportunity to create a student ID to use when submitting surveys at the beginning of the semester when filling out the consent form. Only the investigators had access to the information linking the names to the student IDs. While the instructor had access to this information, being one of the investigators, he did not participate in the distribution or collection of the consent form or the surveys and was not involved in entering the bonus points into the grade sheet, so students could be relatively certain that he would not link their names and responses.

The first questionnaire (pre-test) consisted of demographic questions, such as age, major, number of hours working, etc., as well as questions establishing their history of participating in collaborative group work and their openness to teaching methods involving group work and linking content to real-world events. The questions were primarily five-point Likert scale questions with '1' being low/poor/negative, '3' being neutral and '5' being high/good/positive. A few open-ended questions were included where appropriate. The first questionnaire was distributed the first week of class, time 1 (T1), and collected one week later.

The second questionnaire (post-test) asked students to assess their course experience. It was composed of T1 questions that were modified to reflect that the students had experience with *CIITN* and its teaching methods and additional questions that addressed specific aspects of *CIITN*, which the students would not have been able to answer at T1. The majority of the questions were asked with a five-point Likert scale, '1' being the most negative response, '3' as neutral and '5' as the most positive. The survey was handed out and completed on the last day of class, time 2 (T2), along with the departmental teaching evaluation.

Results

The purpose of this study was to ascertain how effective the *CIITN* intra-group peer review system was at achieving its goals. In order to accomplish this, we collected and analysed survey data regarding students' perceptions of its effectiveness. The means

of students' responses to survey questions were calculated and correlation coefficients were computed for the scores of relevant questions in order to determine how the various perceptions were related. We also examined the system's effect on students' scores. The effect was determined through the computation of means and standard deviations of students' intra-group, portfolio and final project scores. To put these data sets in context, demographic data of the students (age, major, etc.) were collected at the beginning of the semester and their perceptions of group work were collected at the beginning and end of the semester. In addition, enrolment and attrition data were examined to ensure that students were not so dissatisfied that they were dropping the course.

Course characteristics

Organic Chemistry I is a large lecture course with a maximum enrolment of 265. In winter semester 2004, 232 students took the first exam. Of these, 229 remained in the course to participate in *CIITN* intra-group peer review. This is an attrition rate of only 1.2%. Two additional students did not complete the course for personal and academic reasons, for a final course attrition rate of 2.2%. The student scores dataset comprises the scores of the 229 students who participated in *CIITN* intra-group peer review. Of the 229 students that took part in *CIITN*, 214 students completed both the pre- and post-surveys. It is their responses that make up the demographics, perceptions prior to participation and perceptions after participation datasets.

Demographics

These 214 students had a mean age 20.7 years. Within this group, 57% worked either full or part time. Those who were employed reported working an average of 18.4 hours per week (Figure 1). Biology majors accounted for 58.2% of the respondents. Finally, one-third of the group planned to attend medical school.

In order to determine whether students would be pre-disposed to working in groups or not, students were asked at the beginning of the semester if they had participated in group work prior to winter semester 2004. They were also asked several questions about how they felt about group work, whether or not they had participated in group work before this course.

Pre- and post-test perceptions of group work

On the pre-test survey given at T1, 25.5% of the respondents reported having worked in groups before this course (laboratories carried out with a partner were excluded). The students reported that they were positive about group work, as is reflected in a mean of 3.73 on a five-point scale (SD = 1.16). An interesting distinction emerges when the respondents are separated into those who had participated in group activities before this semester and those who had not. Those who had participated in group work before scored the experience at 3.40 (SD = 1.25) while those who had not participated in group work scored it significantly higher at 3.87 (SD = 1.10, $p = .008$).

On the post-test survey, students were asked to rate their overall group experience and if they would take organic chemistry I again, if they had a choice. They were also asked to rate how it influenced their course experience and whether they would choose to work in a group in similar courses in the future. Despite saying they would not take

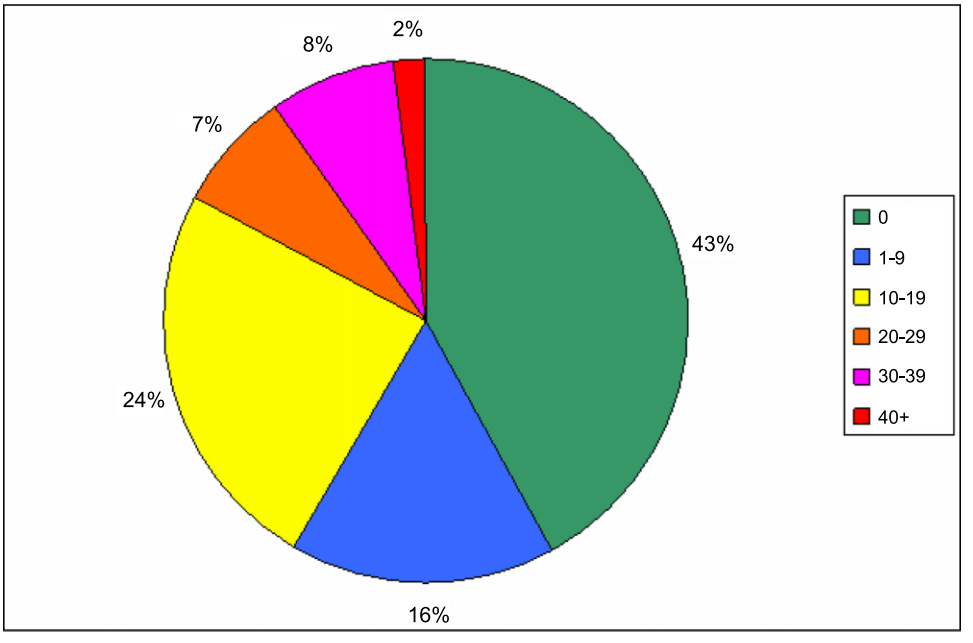

Figure 1. Percentage of employed students by the range of hours worked at job per week.

the course again (2.58, SD = 1.45), students reported post-test perceptions of group work (4.02, SD = 1.07, p = .001) were significantly higher than their pre-test perceptions of group work. They also reported that group work had a positive influence on their course experience (4.09, SD = 1.03). In addition, they were also positive about participating in group work in the future, rating it at 3.92 (SD = 1.36).

In order to gauge change in attitude from the beginning of the semester to the end, the correlation between students' responses to these T2 questions about how they perceived group work and their response to the question at T1 about how they anticipated liking group work. The responses to these questions at T2 were only minimally correlated with their response to the question about how they anticipated liking group work when asked at T1 (Table 1). However, the students who had not participated in

Table 1. Mean perception of group work at T1 and T2.

T2 question	Score (SD)	Correlation with T1[§]
How would you rate your group experience?	4.03 (1.07)	0.31**
How would you rate the affect your group experience had on your overall course experience?	4.09 (1.03)	0.27**
Given the option, would you choose to work in a group in similar courses in the future?	3.92 (1.36)	0.30**

Note: Standard deviation (SD) provided in parentheses.
Strength of Correlation is interpreted as follow: 0.01–0.14= very low; 0.15–0.34 = low; 0.35–0.54 = moderate; 0.55–0.74 = substantial; 0.75–0.94 = very high.

[§]T1 question: How do you like the idea of working in a collaborative student group? (3.73, SD = 1.16)
**Correlation is significant at the 0.01 level (two-tailed).

Table 2. T2 mean perceptions based on prior group experience.

T2 question	Prior experience	No prior experience	*p* value
How would you rate your group experience?	3.79 (1.17)	4.11 (1.04)	0.061
How would you rate the affect your group experience had on your overall course experience?	3.95 (1.11)	4.14 (1.00)	0.219
Given the option, would you choose to work in a group in similar courses in the future?	3.77 (1.48)	3.97 (1.30)	0.345

Note: Standard deviation (SD) provided in parentheses.

any group experiences prior to this semester scored their overall group experience significantly more positively. These students also reported that their group experience's effect on their course experience and their willingness to work in a group in the future were more positive compared to students with prior group experience, though the difference was not significant for these two questions (Table 2).

Intra-group peer review score effect on overall project score

Students are generally very conscious of the process of earning grades throughout the course of the semester. This is particularly true for those students planning on pursuing a medical degree and other post-baccalaureate degrees, and they account for a large portion of the student body in this course. While this may be an issue of concern from the standpoint of students' attitudes about learning (Kohn 1991), assessing how a particular element of the course evaluation affects their final project grade is an important means of both determining the validity of the scoring system as well as part of determining how well students accept the system and a step in explaining why they accept it or not.

The mean intra-group peer review score of the 229 students is 99.22 and the mode is 100 (107 occurrences out of 229 scores); thus most students distributed all of the points and students most frequently earned all 100 points, which is as expected. The interesting outcome of the intra-group peer review scores is not the mean score, but it is instead the range and standard deviation. The range is 3–130 and the standard deviation is 11.63, indicating a fair amount of distinction or variability among students. The comparison between their intra-group scores and the portfolio scores, which averages 91.22 with a standard deviation of only 6.51, is particularly interesting. Because the intra-group and portfolio scores contributed equally to the final score, the intra-group peer review score had a greater effect on the students' final *CIITN* score than the portfolio score (Figure 2).

The effect of the intra-group peer review score is more clearly revealed by examination of the variation of scores within groups. To accomplish this, the standard deviation of intra-group peer review scores was computed for each group. The mean standard deviation for all 49 groups was 6.77 points, though most groups had a standard deviation of 0, with all group members of one-third of the groups receiving 100 points. However, for the remaining two-thirds of the group, substantial differences resulted in the final *CIITN* score. For example, in a group with a standard deviation of 6.75 in intra-group peer review scores, near the class average, the group members had intra-group peer scores of 91, 99, 100, 100 and 110. This produced final scores ranging

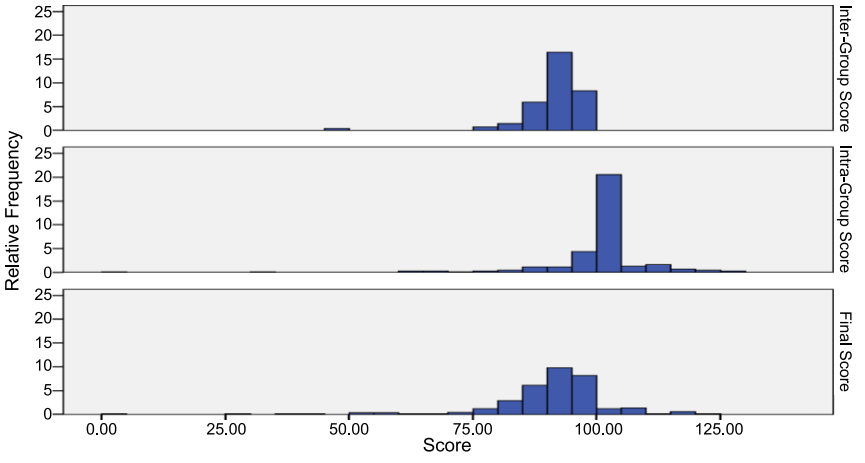

Figure 2. Distribution of *CIITN* Scores. The intra-group peer-review score is converted to a percentage and multiplied by the portfolio's inter-group score to arrive at the final score.

from 81.9 to 99 when these intra-group portfolio scores were converted to percentages and multiplied by the portfolio score of 90. The group with the largest standard deviation (SD = 54.23) had intra-group peer review scores of 3, 124, 124, 124 and 125. This resulted in final scores ranging from 2.8 to 117.1 with a portfolio score 93.7 (Figure 3).

As collaboration is seen as integral to the process of portfolio construction as opposed to a convenient means of having students complete such a process, we were interested in determining if the portfolio scores were related to the intra-group peer review scores. It was found that the portfolio score was not correlated with the intra-group peer review score (r = .06). The relationship between standard deviation of intra-group peer review scores within groups and the portfolio scores were examined, and it was discovered that variability of scores within groups was also not correlated with the group score (r = –.07). Finally, the number of hours a student worked (to earn money), which one might expect to hinder collaboration, was not significantly correlated with his/her intra-group peer review score and minimally negatively correlated with his/her final score (r = –.23, p = .011). However, the average number of hours worked by all students within the group was moderately negatively correlated (r = –.40, p = .004) with the group's portfolio score.

Students' perception of the intra-group peer review system

How students perceive the system is particularly important in light of how the intra-group peer review affects final scores. At T1, students were asked if they thought collaborative group work would enhance their learning, improve their grade and increase their interest. They responded moderately positively with scores of 3.90 (SD = 1.04), 3.80 (SD = 1.07) and 3.64 (SD = 1.09), respectively. At T2, students were asked how the group experience affected their course performance and they also responded positively (4.11, SD = 0.98). Based on these responses we expected students would be more likely to perceive the system positively.

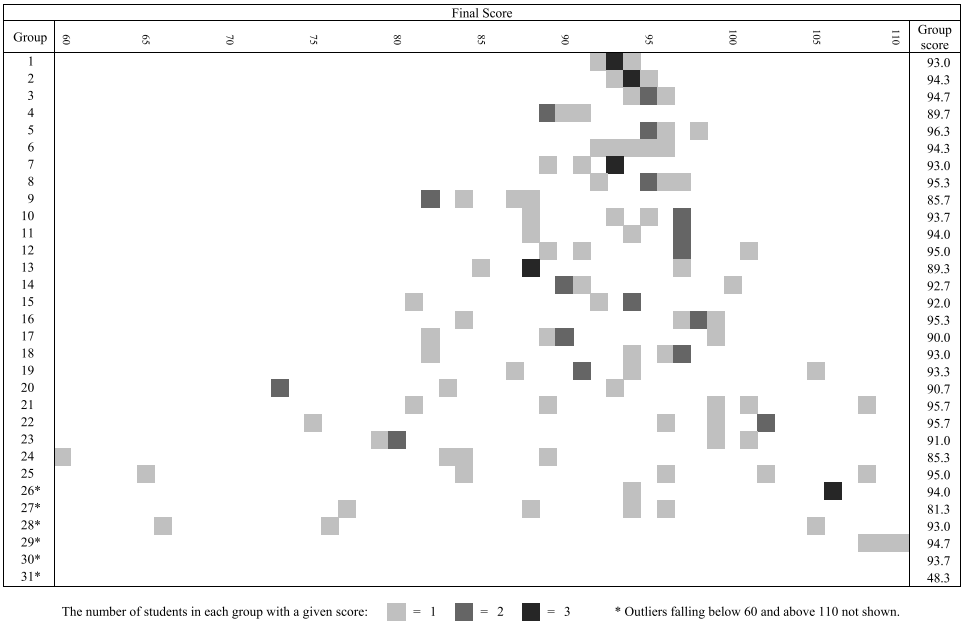

Figure 3. Distribution of final scores within groups. Example of the use of the information provided: The five students in Group 6 obtained different final scores and each one of their scores of 92, 93, 94, 95 and 96 is indicated by a lightly shaded square. Three of the five students of Group 1 obtained the same score of 93 and this is indicated by the darkest square. The final scores of the other two students in Group 1 are 92 and 94, respectively, and the single occurrences of these scores are indicated again by lightly shaded squares.

At T2, students were asked how they would rate the intra-group peer review system overall and they responded positively, rating it at 4.03 (SD = 1.10). More specifically, students were asked whether they perceived the intra-group peer review to be fair and accurate. The students also responded positively to these questions, rating them 4.12 (SD = 1.16) and 4.33 (SD = 1.06), respectively. The scores for every one of these three questions are above 4.00 and this result is very encouraging.

We were also interested in determining if the students' perceptions about group work before and after the semester are correlated with the overall perception of intra-group peer review as well as the perception of its fairness and accuracy. While the students' perception about group work at T1 was only minimally correlated with the students' perception of the system overall, its fairness and its accuracy, students' perception of group work at T2 was moderately to substantially correlated with all three (Table 3).

In addition to the positive scores for fairness and accuracy, there was only one appeal to the teaching staff about a student's intra-group peer review score. This is remarkable considering how frequently students appeal exam and other scores. The teaching staff compared the student's individual score, examined the comments and discussed the observations of the situation. It was determined that the intra-group peer review system had been abused by one of the group's members, and the decision was made to suspend the student's intra-group peer review score and to award her the group's portfolio score as a final score. No other appeals were submitted and the low

Table 3. Correlation between T1 and T2 perceptions of peer review.

T2 question	T1 perception correlation	T2 perception correlation
Rate your *CIITN* intra-group assessment experience.	0.26**	0.53**
Rate the *CIITN* intra-group assessment fairness.	0.24**	0.56**
Do you feel like your group members' assessment of you is accurate?	0.12	0.48**

**Correlation is significant at the 0.01 level (two-tailed).

scores, for example 3 out of 100, were supported by intra-group peer review comments such as 'I never met him'.

Student experience using intra-group peer review

Finally, it is important to determine how students felt about engaging in intra-group peer review. To establish this, students were asked if they felt comfortable or qualified assessing their group members, if they would feel comfortable or qualified assessing peers in the future and how much peer pressure affected their scores. Students responded even more positively to the first two questions than they had to the questions about the system overall and its fairness, scoring them 4.36 (SD = 0.97) and 4.22 (SD = 1.01), respectively. In addition, students also reported that peer pressure had little effect on their scores, rating its effect at 2.44 (SD = 1.43).

We wanted to establish how this was related to students' perceptions towards group work at T1 and T2 as well as how they felt about the peer review system overall, its fairness and its accuracy (Table 4). The responses to the questions about being comfortable engaging in intra-group peer review now and in the future were minimally to moderately correlated to the responses to questions regarding perceptions about group work. The response to how peer review affected intra-group peer review scores, however, had a very low to low correlation with responses about perceptions group work.

Table 4. Correlation between perception and comfort of students using peer review.

Question	Comfortable	Comfortable in the future	Peer pressure
How do you like the idea of working in a collaborative student group? (T1)	0.19**	0.20**	–0.09
How would you rate your group experience? (T2)	0.20**	0.22**	–0.12
Rate your *CIITN* intra-group assessment experience. (T2)	0.39**	0.34**	–0.23**
Rate the *CIITN* intra-group assessment fairness. (T2)	0.38**	0.35**	–0.15*
Do you feel like your group members' assessment of you is accurate? (T2)	0.31**	0.26**	–0.12

*Correlation is significant at the 0.05 level (two-tailed).
**Correlation is significant at the 0.01 level (two-tailed).

Discussion

Overall, the *CIITN* intra-group peer review system fulfiled the goals set forth. It provided student differentiation as evidenced by the standard deviation of the intra-group peer review scores. The mode remained 100. This may appear negative in that it might be taken as indicating limited differentiation in scores and possible grade inflation, but that is not necessarily the case. The hope is, obviously, that groups work well together and if they do, it is appropriate that students score 100. Thus a mode of 100 is positive as long as it corresponds with students' feelings that the system is fair and accurate, which was the case in the semester under examination. In addition, there were a substantial number of students who received scores other than 100 as well as a large standard deviation of scores, indicating ample differentiation. Moreover, the intra-group peer review system avoided inflating grades through two means. First, because awarding one student with more than 100 points means scoring another student lower, students are encouraged to score moderately. Second, because the intra-group peer review score is converted to a percentage score and multiplied by the portfolio score, the final score is still linked to the quality of the product of their collaboration.

Demographics

The demographics of the students in this course illustrate that it is populated primarily by traditional students as far as their age is concerned. However, the large percentage that work, particularly the very discouraging proportion of students who work more than 20 hours per week, mean that this course cannot be assumed to operate as a class of traditional students. These students have significant demands on their time aside from their course work and, consequently, scheduling time for group work is a substantial challenge for everyone in the course. The findings of the correlations between work hours and lower portfolio scores for the group as well as previous findings, in which students complained that scheduling was a significant problem with group work (Hume et al. 2006), supports these findings. Therefore, positive responses about group experience at the end of the semester are even more persuasive because these students are not easily able to accommodate group activities in their schedules.

In addition to the limited outside-of-class time students have to devote to collaborative group activities, other student characteristics need to be taken into account. First, the large number of biology majors, greater than half of the respondents, means that these students, by and large, are not motivated to take this course out of interest, as one might assume if they were chemistry majors. Instead, organic chemistry I is a requirement, one which most students would not take (again) if they were given a choice. Second, the large number of students who plan to attend medical school also has implications as to how they approach this course. Students wanting to apply for medical school are aware they need a good grade point average and, in particular, need good grades in required courses such as organic chemistry courses. Both of these characteristics do not bode well for group work in such a course. The first means that students are not going to be particularly amenable to spending the 'extra' time on this course that collaborative group work might entail, and the second means that they are going to be particularly anxious about their grade and therefore less comfortable trusting it to fellow students by engaging in collaboration, which is anecdotally supported by student commentary. In addition, because applying to medical

school is a highly competitive process, students on the pre-med track have the tendency to see every pre-med course as a venue in which to prove themselves a superior candidate.

Group work overall

Despite the demographics data that would lead one to assume that the students would not want to participate in collaborative group work, students anticipated liking collaborative group work. The finding that there is a significant difference between those that have participated in group work before this course and those that have not at T1, shows that students find the collaborative group activity more appealing in theory than in practice. However, the significant improvement in the score of the perception of group work at T2 means that the *CIITN* group experience appears to have helped students rebound from less positive group work experiences prior to this semester and to have started the remaining students out on the right foot for collaborative group work. This is also manifested in the significant difference in the two groups, those with prior experience and those without when asked about their overall group experience at T2. This finding is noteworthy because it points to the need for a supportive environment for collaborative group work. Since creating and maintaining such an environment was a primary consideration in the design of the intra-group peer system, it can be judged as having succeeded.

Intra-group peer review score effect on overall project score

As stated earlier, the intra-group peer review system had the desired affect of producing differentiation among students, and therefore, its effect on the overall project score is also as desired: a greater distribution of scores. The effect was not perfectly applied, however. Ideally, the group would recognise that a lower-quality portfolio is usually a product of lower-quality collaboration and therefore, the scores would be lower in that group but there is no correlation between the portfolio score and the scores that the group members assign each other in intra-group peer review. There are two possible explanations. The first is that students who are less invested in the production of the portfolio are also going to be less invested in discriminating the quality of contribution by their group members, and thus score members more generously than would be merited. The second is simply that it is an artefact of the system itself. Though students have the option of not assigning all points, students still prefer to assign all the points. This is not a problem when it comes to validity of final scores because of the Goldfinch and Raeside (1990) system adaptation.

In addition to meeting the scoring goals, despite the lack of correlation between the two scores, students felt that the system was both fair and accurate. Thus, the system met the initial motivation driving the implementation of an intra-group peer review system: improving student satisfaction with the assessment of the *CIITN* project. This result also indicates that the system produces valid results as the intra-group peer review system is designed to measure how well students meet their fellow students' expectations. A positive rating for both fairness and accuracy means that students formulated and communicated these expectations well and felt that these expectations were reasonable.

The report that peer pressure has a relatively little impact on the students' scoring suggests that the safety mechanisms built into the intra-group peer review system are

effective. In addition, students felt comfortable using the system and would feel comfortable assessing their peers in the future and these findings indicate that the system is both well designed and students understand its elements and expectations.

Limitations

The primary limitation of this study is that there is no formalised assessment or analysis of how the intra-group peer review system was presented to the students. This is a somewhat difficult proposition to accomplish in retrospect when the bulk of the presentation of the system is done in computer training sessions by three different TAs whose conceptualisation and presentation of both the purpose and function of intra-group peer review may vary.

Implications

The findings show that the system was effective at achieving its goals; however, there is room for improvement. The primary area for improvement for the system would be in its framing for students. Better communication of the system's goals would be evident in larger variability as well a better correlation between the intra-group peer review and the portfolio scores.

There is also room for further improvement in the support for collaborative group work, which would also improve the students' use of the intra-group peer review system. To accomplish this, we will integrate some elements of Peer-Led Team Learning model (Gosser et al. 2001; Gosser and Roth 1998) into the course. It is hoped that this will further help students to productively engage in collaborative group work.

Future research

The intra-group peer review system will be re-assessed in light of these changes. In particular, attention will be paid to how the intra-group peer review system is presented to students. This in combination with an even more supportive environment for collaborative group work which will help improve the psychometric properties of the intra-group peer review system and thus collaborative group work assessment.

Acknowledgements

We thank Dr Jennifer Hart for her constructive review of the thesis on which this paper is based. We also thank the MU Institute for Instructional Development (IID), the Camille and Henry Dreyfus Foundation and the National Science Foundation for support.

Notes on contributors

Kathleen M. Carson received a BA in International Studies at MU in 2003 and an MA in Educational Leadership and Policy Analysis at MU in 2005. She has focussed on interdisciplinarity in the pursuit of her education, having a background in both hard science and social science and combining her graduate classroom experience in the College of Education with research experience in the Department of Chemistry. She is the recipient of the 2005 MU Distinguished Master's Thesis Award. She is currently pursuing a PhD in Public Policy at St Louis University.

Dr Rainer E. Glaser, professor of chemistry, studied chemistry and physics in Tübingen, at Berkeley, and at Yale. He is a physical organic chemist with three major lines of research in bioorganic chemistry and toxicology, materials chemistry and astrochemistry. He has collaborated extensively with chemists, biochemists, physicists, mathematicians, astronomers, educators and journalists. In 1995, Glaser began his education research with the novel curriculum, *Chemistry Is in the News* (*CIITN*), that he designed for chemistry education of science majors. He was a JSPS Fellow in 1997 and was elected AAAS Fellow in 2004 and Fellow of the Royal Chemical Society in 2006.

References

Ballantyne, R., K. Hughes, and A. Mylonas. 2002. Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education* 27, no. 5: 427–41.

Barfield, R.L. 2003. Students' perceptions of and satisfaction with group grades and the group experience in the college classroom. *Assessment & Evaluation in Higher Education* 28, no. 4: 355–69.

Boud, D., R. Cohen, and J. Sampson. 1999. Peer learning and assessment. *Assessment & Evaluation in Higher Education* 24, no. 4: 413–26.

Bransford, J.D., A.L. Brown, and R.R. Cocking. 2000. *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Carlone, H.B. 2004. The cultural production of science in reform-based physics: Girls' access, participation, and resistance. *Journal of Research in Science Teaching* 41, no. 4: 392–414.

Carson, K.M., B. Hodgen, and R.E. Glaser. 2006. Teaching dissent and persuasion. *Education Research and Reviews* 1, no. 4: 115–20.

Carson, K.M., D.L. Hume, Y. Sui, S. Schelble, and R.E. Glaser. 2009. *Chemistry Is in the news*: The why and wherefore of integrating popular news media into the chemistry classroom. In *Chemists' guide to effective teaching,* Vol. 2, ed. T.J. Greenbowe, M.M. Cooper, and N.J. Pienta, 230–45. Prentice Hall Series in Educational Innovation. Upper Saddle River, NJ: Prentice Hall.

Colburn, A. 2000. Constructivism: Science education's 'grand unifying theory'. *The Clearing House* 74, no. 1: 9–12.

Conway, R., and D. Kember. 1993. Peer assessment of an individual's contribution to a group project. *Assessment & Evaluation in Higher Education* 18, no. 1: 45–57.

Cooper, M. 2005. An introduction to small-group learning. In *Chemist's guide to effective teaching,* ed. N.J. Pienta, M.M. Cooper, and T.J. Greenbowe, 117–28. Upper Saddle River, NJ: Prentice Hall.

Fehr, E., and S. Gächter. 2000. Cooperation and punishment in public goods experiments. *The American Economic Review* 90, no. 4: 980–94.

Fehr, E., and S. Gächter. 2002. Altruistic punishment in humans. *Nature* 415: 137–40.

Glaser, R.E., and K.M. Carson. 2005. Chemistry Is in the news. Taxonomy of authentic news media based learning activities. *International Journal of Science Education* 27: 1083–98.

Glaser, R.E., and M.J. Poole. 1999. Organic chemistry online: Building collaborative learning communities through electronic communication tools. *Journal of Chemical Education* 76: 699–703.

Glaser, R.E. 2003. Science communication for all. *Chemistry International* 23: 3–6.

Glaser, R.E., Z. Wu, and Y. Sui. 2004. Chemistry Is in the news [webtool]. http://ciitn.missouri.edu (accessed April 30, 2004).

Goldfinch, J. 1994. Further developments in peer assessment of group projects. *Assessment & Evaluation in Higher Education* 19, no. 1: 29–36.

Goldfinch, J., and R. Raeside. 1990. Development of peer assessment technique for obtaining individual marks on a group project. *Assessment & Evaluation in Higher Education* 15, no. 3: 210–31.

Gosser, D.K., M.S. Cracolice, J.A. Kampmeier, V. Roth, V.S. Strozak, and P. Varma-Nelson. 2001. *Peer-led team learning – A guidebook.* Upper Saddle River, NJ: Prentice-Hall.

Gosser, D.K. Jr., and V. Roth. 1998. The workshop chemistry project: Peer-led team-learning. *Journal of Chemical Education* 75: 185–7.

Gupta, M.L. 2004. Enhancing student performance through cooperative learning in physical sciences. *Assessment & Evaluation in Higher Education* 29, no. 1: 63–73.
Hall, E.W. 2003. Chocolate chip cookies and rubrics: Helping students understand rubrics in inclusive settings. *Teaching Exceptional Children* 35, no. 4: 8–11.
Hume, D.L., K.M. Carson, B. Hodgen, and R.E. Glaser. 2006. Chemistry Is in the news. Philosophy, pedagogy and assessment of authentic news media based learning activities. *Journal of Chemical Education* 83: 662–7.
Kohn, A. 1991. Group grade grubbing versus cooperative learning. *Educational Leadership* 48, no. 5: 83–7.
Kruck, S.E., and H.L. Reif. 2001. Assessing individual student performance in collaborative projects: A case study. *Information Technology, Learning, and Performance Journal* 19, no. 2: 37–47.
Lotan, R.A. 2003. Group-worthy tasks. *Educational Leadership* 6, no. 6: 72–5.
Marsh, H.W., U.W. Jayasinghe, and N.W. Bond. 2008. Improving the peer-review process for grant applications: Reliability, validity, bias, and generalizability. *American Psychologist* 63, no. 3: 160–8.
McKeachie, W.J. 2002. Active learning: Cooperative, collaborative, and peer learning. In *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers* (11th edn), ed. W.J. McKeachie, 187–95. Boston: Houghton Mifflin.
Olson, M. 1982. *The rise and decline of nations.* New Haven, CT: Yale University Press.
Phillip, C. 2002. Clear expectations: Rubrics and scoring guides. *Knowledge Quest* 31, no. 2: 26–7.
Searby, M., and T. Ewers. 1997. An evaluation of the use of peer assessment in higher education: A case study in the School of Music, Kingston University. *Assessment & Evaluation in Higher Education* 22, no. 4: 371–83.
Seymour, E. 2002. Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology. *Science Education* 86, no. 1: 79–105.
Seymour, E., and N. Hewitt. 1997. *Talking about leaving: Why undergraduates leave the sciences.* Boulder, CO: Westview Press.
Slavin, R.E. 1996. Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology* 21: 43–69.
Smist, J. 2004. Small groups are worth the effort. In *Survival handbook for the new chemistry instructor,* ed. D.M. Bunce and C.M. Muzzi, 97–104. Upper Saddle River, NJ: Prentice Hall.
Strom, R.D., and P.S. Strom. 1998. Student participation in the evaluation of cooperative learning. *Community College Journal of Research and Practice* 22, no. 3: 265–79.
Tinto, V. 1997. Classrooms as communities: Exploring the educational character of student persistence. *Journal of Higher Education* 68: 599–623.
Woolfolk, A. 2004. *Educational psychology,* 9th ed. Boston: Pearson.
Wu, Z., and R.E. Glaser. 2004. Software for the synergistic integration of science with ICT education. *Journal of Information Technology Education* 3: 325–39.
Yager, R.E., and Lutz, M.V. 1994. Integrated science: The importance of 'how' versus 'what'. *School Science and Mathematics* 94, no. 7: 338–47.

Appendix 1. Intra-group peer review rubric

Level of performance	Highly effective	Effective	Moderately effective	Not effective	Inadequate
	Truly went above and beyond in an effort to improve the group and the group's project	Contributed their share to the group, fulfiling what was required of them and enhancing the group and final project	Attempted to participate but occasionally fell below the level expected of them, overall moved the group forwards	Frequently fell below the level expected or did not attempt the role but was not severely disruptive to the group or the project	Disruptive in their attempt to fulfil the role or their refusal to take on the role, detrimental to group cohesiveness and overall quality of final project
Facilitator Proposer Supporter Critic Organiser	Attempted all five roles	Attempted four out of five roles	Attempted three out of five roles	Attempted one–two out of five roles	Attempted no roles
Time contribution	Attended all meetings and did significant outside research	Attended all meetings and did some outside research	Attended most meetings and did little outside research	Attended most meetings but did no outside research	Attended few meetings and did no outside research
Share of workload	Picked up any 'slack'	Fulfiled expectations completely	Fulfiled most expectations	Fulfiled a minimum amount of expectations	Did not fulfil expectations
Project knowledge	Exceptional awareness of project requirements and how to meet them, innovative ideas provided	Aware of project requirements and how to meet them creatively	Mostly aware of project requirements and how to meet them conventionally	Generally aware of the project requirements but is unsure of how to meet some of them	Unaware of most or all of the project requirements and few or no ideas of how to meet them
Chemistry knowledge	Exceptional chemistry knowledge provided to the group	Provided group with the knowledge of chemistry needed	Adequate chemistry knowledge shared with group	Inadequate knowledge of chemistry shared with group	Share no chemistry knowledge with group

Appendix 1. (*Continued*)

Level of performance	Highly effective	Effective	Moderately effective	Not effective	Inadequate
Ability to cooperate	Enhanced cooperation with the group as a whole	Cooperated as expected	Cooperated as needed	Cooperated at times but not on a regular basis	Did not cooperate to any meaningful extent
Enthusiasm for group work	Highly enthusiastic	Showed overall enthusiasm	Showed some enthusiasm	Was primarily unenthusiastic	Displayed a negative attitude
Similarity of goals for project with group	Complete agreement with group's goals	Primarily in agreement with group – worked towards group's goals	Adequate amount of agreement 0 worked towards most goals	Inadequate amount of agreement – worked towards few goals	Total inadequate amount of agreement – only concerned with personal goals
This person's contribution to this project made it	Significantly better	Moderately better	No better or worse	Moderately worse	Significantly worse