Peer and Self-evaluation: Individual Accountability in Teams

KIM L WILSON

INTRODUCTION

AS TEAM-BASED PROJECTS become more popular in university landscape architecture programmes, effective ways to evaluate students should be implemented. In the first two years of using semester-long teams in a senior urban design studio both students and instructor felt that team grades were not a fair reflection of individual work and the instructor was uncomfortable with altering individual grades based solely on class observation because a significant portion of the work was completed outside the classroom. Furthermore, it was difficult to differentiate grades among teams, resulting in a narrow grade range. Because team projects produced quality work, grades also were uniformly high, where 92 percent of the students received A's.

Effective evaluations include mechanisms for measuring individual contributions and overall team performance. A combination of product evaluation by the instructor, peer evaluation by team members, and self-evaluation by each student is necessary to obtain a comprehensive summative evaluation (Crews and North, 2000; Falchikov, 1991). Implementing self and peer evaluations requires instructors to build a solid foundation for incorporating self and peer evaluation into the classroom as well as develop evaluation criteria, ensure honest student participation, and implement both formative and summative feedback (Michaelsen et al, 2002). This study tested a method that uses a combination of product evaluation by instructor, peer evaluation by team members and self-evaluation by each student. The study tested:

- The reliability and fairness of peer and self-evaluation in identifying the students who look to shirk their responsibility (social loafers or hitchhikers) (Slavin, 1985; Levi and Cadiz, 1998).
- The reliability of peer and self-evaluation versus instructor observation.
- The reliability of formative evaluation versus summative evaluation.

This paper describes the methodology used for peer and self-evaluation in teams, including the studio process, instrument, and grading model based on best practices. Results of and recommendations on peer and self-evaluation in teams are presented.

METHOD

The peer and self-evaluation method was administered to semester-long teams, in two landscape architecture senior urban design studios, in two consecutive years. A total of 52 students were included in 12, four-to-five member teams.

Kim L Wilson, Associate Professor, Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, Indiana 47907, United States of America. Email: wilson@hort.purdue.edu

KEY WORDS

Peer evaluation
Self-evaluation
Individual accountability
Teams
Criterion-referenced evaluation
Formative evaluation

Process

The four-step process included building a foundation for the evaluation, establishing evaluation criteria, administrating the evaluation and supporting student learning. The following describes each step.

First, a foundation was laid for student-centred learning by providing the student with a clear presentation of who would evaluate them, what the evaluation included, when the evaluation would be done, and how the peer and self-evaluation affected their grade (Webb, 1993). The most effective way to alleviate student concerns about team evaluations is to involve students in the development of the grading system, where students adopt a degree of responsibility and ownership of the evaluation process (Michaelsen, 2002; Ramsden, 1992).

Secondly, students developed the peer evaluation criteria (Johnson, 1993; Michaelsen, 2002). This involved creating awareness of individual behaviours required for team performance through a team-building workshop where the acceptable behaviours were identified, defined, agreed upon, and translated into a 'contract' to be used as the peer and self-evaluation. An in-class training session was conducted to instruct students how to evaluate each other using the accepted criteria (Ramsden, 1992).

Thirdly, peer and self-evaluations were administered by the instructor after the completion of the three phases of the semester project. Evaluation forms were completed and returned in the last 15 minutes of the class to ensure accurate student input and total class participation. After the last phase of the project, the instructor evaluated each student based on classroom observation using the same instrument. In addition to the category-based criteria evaluation, open-ended questions or formative assessment were included in every evaluation for qualitative feedback to questions concerning the team process. Students also made weekly journal entries about the project and team. Each journal was reviewed twice during the semester and at the end of the semester.

Lastly, one week after each evaluation, the instructor compiled the assessment data and shared the information with team members as a team building exercise. When the peer and self-evaluations, assessment, or weekly journal entries identified issues affecting team performance, the instructor recommended a team meeting with a trained mediator. The mediator knew in advance the issues and the backgrounds of team members to assist in the team-building process.

Instrument

The peer and self-evaluation criteria and format used in this study were developed by the University of Illinois, Champaign-Urbana Campus, Center for Instructional Excellence, for peer evaluation in their MBA programme (see appendix).

The six behaviour criteria used addressed individual contribution in teams and included attendance at team meetings, timeliness, collaboration, effort, contribution of skills and resources, and contribution of ideas. Students were introduced to these behaviours as part of a team skill-building exercise where the six behaviours were presented and discussed by a trained presenter. Teams were then asked to continue their discussion of the criteria outside of class and add new

criteria and/or alter any definitions based on team consensus. The criteria were formalised in a handout, with definitions, and teams were given one last opportunity to change any of the definitions. Upon acceptance by all teams, the criteria and definitions were used for peer and self-evaluation for the duration of the semester-long project. Each peer and self-evaluation criteria was assigned a five-point Likert scale from almost never (1) to almost always (5) resulting in a possible combined total for all six criteria of 30 points (see appendix).

Formative assessment questions were listed after the evaluation criteria for all three phases. The final evaluation asked each student to provide a summative evaluation of individual contribution of each team member and themselves by assigning a letter grade (A through F), and a final grade for the team project and team.

Grading model

A criterion-referenced evaluation grading scheme was used where absolute criterion is set (90 percent for an A) and anyone and everyone who meets or exceeds this criterion receives that grade (all or no-one can earn an A) (Johnson, Johnson, and Smith 1998; Smith, 1998). The students' final grade was the total accumulated points for individual contribution based on peer and self-evaluation, plus team product points assigned by the instructor with 1,000 points possible.

The peer and self-evaluation contributed 300 possible points of 1,000 total possible points in this class, or 30 percent of the final grade. The recommended percentage of peer and self-evaluation range from 10 to 30 percent of the final grade, and higher percentages are used in project-based courses, such as senior design projects in engineering (Michaelsen, 2002; Smith, 1998).

The second peer and self-evaluations were weighted two times, and the final-three times that of the first evaluations. The total possible points for each phase: phase I – 50 points; phase II – 100 points; and phase III – 150 points = 300 possible points.

Because a team project typically involves a series of steps or phases, team dynamics change throughout the progression of the team experience. Team development goes through fairly predictable stages of forming, storming, norming and performing. Increasing the weight of the peer and self-evaluations over the three phases encourages students to develop team skills early in the collaborative process. More weight is placed on the last two phases of the project where conflict would likely arise because of the anxiety generated when teams choose between or among alternatives and disagreements over procedures needed to guide the team to project completion (Engolf, 2001).

The instructor evaluated the four required team products where all team members received the same grade. The combined total of team products was 600 possible points.

RESULTS AND DISCUSSION

Eleven out of the 12 teams received an A (92 percent of all students) for team performance. Individual grades ranged from 28 A's (54 percent); 19 B's (37 percent); and 5 C's (9 percent). Peer and self-evaluation lowered 39 percent of the final grades to B's and C's. Because the journal entries verified that the peer self-

evaluations were not biased, these evaluations had a significant effect on individual grades. Therefore, we can conclude that peer and self-evaluations are effective.

The reliability of peer and self-evaluation versus instructor observation.
The results showed that 61 percent of peer and self-evaluations were the same as the instructor's observations. Six percent of the instructor's observations were higher than the peer and self-evaluations, where as 29 percent were lower. The differences may be for a variety of reasons including students' grade themselves higher, expectations, and instructor's limited exposure to team interaction. However, the results verify that, when substantial out-of-class collaboration is required, students are found to be more accurate evaluators of the relative contributions of the other team members (Johnson, 1993).

The reliability of formative evaluation versus summative evaluation.
There was a 36 percent difference between the summative and formative aspect of the peer and self-evaluations, with 32 percent of the summative evaluations higher than the total formative evaluation. The phase III summative evaluation was 30 percent higher than the formative evaluation, even though they were taken on the same day. These results indicate that the specificity of the evaluation criteria produces better quantitative differentiation of grades and may be more accurate than summative evaluation. If peer evaluation is left until the end of the team project (summative), students are not able to re-direct themselves or the team towards a more successful approach during the team experience.

CONCLUSION

Peer and self-evaluation of individuals on teams is different and independent from their team project grades. Allowing students' evaluation of team work is an important part of the grading system. Team work scores discriminate within team performance in ways that should be included in the overall individual grades. In addition, the value of peer and self-evaluation improves the accuracy of grading, allows students to manage social loafing in their teams, and gives the teams a way to learn and improve. Self and peer evaluation should be part of a process towards student-centred learning. The change requires a shift in emphasis from the norm-reference to criterion-referenced, from purely summative to formative plus summative, and from evaluation of product to evaluation of process.

REFERENCES

Crews, TB and North, AB (2000) Team evaluation (Part 2 of 2), *Instructional Strategies*, vol 16, no 2, pp 1-4.

Engolf, DB (2001) *Forming, Storming, Norming, Performing: Successful communication in groups and teams*, New York: Writers Club Press.

Falchikov, N (1991) Group process analysis: self and peer assessment of working together in a group, *Standing Conference on Educational Development*, SCED, Paper 63.

Johnson, DW; Johnson, RT and Smith, KA (1998) *Active Learning: Cooperation in the college classroom*, Edina, Minnesota: Interaction Book Company.

APPENDIX 1: PEER AND SELF-EVALUATION INSTRUMENT

Evaluation criteria

a. **Attendance at group meetings:** present at all team meetings except where a previous commitment conflicted with the time and the absence was agreed upon with team members *in advance*.

b. **Timeliness:** completed all assigned tasks in a timely manner.

c. **Collaboration:** made a genuine effort to work effectively with others.

d. **Effort:** exhibited a high level of interest and commitment to the assignment.

e. **Contribution of skills and resources:** made available valuable competencies, resources and materials.

f. **Contribution of ideas:** provided creative and innovative ideas for group discussion.

Almost Always	Frequently	Sometimes	Infrequently	Almost Never
5	4	3	2	1

Your name or team member's name: _____

Attributes:	Almost always				Almost never
a. Attendance at group meetings	5	4	3	2	1
b. Timeliness	5	4	3	2	1
c. Collaboration	5	4	3	2	1
d. Effort	5	4	3	2	1
e. Contribution of skills and resources	5	4	3	2	1
f. Contribution of ideas	5	4	3	2	1

Johnson, KF (1993) Team peer evaluations: A student-generated quantitative measurement of group membership performance, paper presented at the *Annual Meeting of the Association for Education in Journalism and Mass Communication*, 76th, Kansas City, Missouri.

Levi, D and Cadiz, D (1998) Evaluating teamwork on student projects: The use of behaviorally anchored scales to evaluate student performance. ERIC Document Reproduction Service, ED 424250.

Michaelsen, LK; Knight, AB and Fink, LD (2002) *Team-based Learning: A trans-formative use of small groups,* Connecticut: Praeger Publishers.

Ramsden, P (1992) *Learning to Teach in Higher Education*, London: Routledge.

Smith, KA (1998) Grading cooperative projects, *New Directions for Teaching and Learning*, Summer, 74, pp 69–73.

Slavin R (1985) *Learning to Cooperate, Cooperating to Learn*, New York: Plenum Press.

Stefani, LAJ (1994) Peer, self and tutor assessment: relative reliabilities, *Studies in Higher Education*, vol 19, no 1, pp 69–75.

Webb, NM (1993) *Group collaboration in assessment: Competing objectives, processes, and outcomes, Project 2.1: Designs for assessing individual and group problem solving: Effects of group characteristics on groups and individual performance*, National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, California: ERIC Documentation Reproduction Services, ED 376215.