The Effect of Scoring Criteria Specificity on Peer and Self-assessment

PETER J. MILLER, *University of the Sciences in Philadelphia, Philadelphia, Pennsylvania, USA*

ABSTRACT   *The value of peer and self-assessments is commonly diminished by scoring range restriction by the raters. This investigation studied the effect of different levels of scoring criteria specificity and written feedback requests on the distribution of scores, the correlation between peer and self-assessments and the quantity and nature of written feedback. Increasing the number of criteria decreased the mean scores and increased the standard deviations of the peer and self-assessments, providing a wider range of scores and increasing the sensitivity of the instrument. Correlation between peer and self-assessment was improved with more specific criteria, depending on the statistic used. However, analysis revealed that the more specific written feedback requests elicited more peer feedback. Educators should consider the effects of criteria specificity and written feedback solicitation on rater behaviour when designing these instruments.*

Introduction

Peer assessment and self-assessment of student performance in small group work are becoming more common in higher education, because they can be used for both formative and summative purposes (Orsmond *et al.*, 1996; Das *et al.*, 1998; Topping, 1998; Purchase, 2000; Li, 2001) and because peer and self-assessment are recognized as skills necessary for professional practice (DeMarchais & Vu, 1996; Heylings & Stefani, 1997; Pond & ul-Haq, 1997; Thomas, 1997; Das *et al.*, 1998; Sullivan *et al.*, 1999; Sluijsmans *et al.*, 2001).

Peer and self-assessment are sometimes incorporated as part of a triangulated approach to assessment, in which student learning is evaluated from multiple data sources or multiple assessors (Denzin, 1989; Breitmeyer *et al.*, 1993; Denzin & Lincoln, 1994; Bailey, 1997). A major issue to be considered when employing triangulation is whether the purpose of it is to achieve convergence, i.e. agreement among the assessment sources, or whether it is to achieve completeness, the uncovering of multiple perspectives on the behaviour being assessed (Breitmeyer *et al.*, 1993; Sim & Sharp, 1998). Studies dealing with peer and self-assessments have commonly concentrated on

the agreement between them and with faculty assessments (Falchikov & Goldfinch, 2000; Magin, 2001).

Peer and self-assessment instruments have commonly consisted of 4–6 criteria for scoring, with the nature of the criteria depending on whether the assessment was directed at learning outcomes (such as the quality of an oral presentation, written paper or other product) or learning processes (such as the level of student participation in a group project) (Orsmond *et al.*, 1996, 2000; Heylings & Stefani, 1997; Lopez-Real & Chan, 1999; MacAlpine, 1999). Rating scales for these assessments have usually involved the assignment of a numerical score for each item in a set of criteria, with 4–5 point scales being common (Conway *et al.*, 1993; Orsmond *et al.*, 1996, 2000; Culbertson *et al.*, 1997; Heylings & Stefani, 1997; Longhurst & Norton, 1997; Pond & ul-Haq, 1997; Cheng & Warren, 1999; Sullivan *et al.*, 1999; Li, 2001). The criteria by which different scores are given are described in detail in some studies, while others are unexplained.

A problem sometimes reported with the use of peer and self-assessments has been the tendency of raters to assign a very narrow range of scores, usually at the high end of the rating scale (Conway *et al.*, 1993; Falchichov, 1995; Das *et al.*, 1998; Cheng & Warren, 1999; MacAlpine, 1999; MacPherson, 1999; Miller, 1999; Pond & ul-Haq, 1999; Sullivan *et al.*, 1999; Purchase, 2000). This tendency is frequently related to issues concerning the use of students as assessors, including their lack of ability to discriminate levels of performance and their reluctance to judge their peers or themselves (Falchikov, 1995; Orsmond *et al.*, 1996; Sullivan *et al.*, 1999; Li, 2001; Sluijsmans *et al.*, 2001). However, it should be noted that the same scoring tendencies have been seen with faculty assessors (Miller, 1999).

The problem with scoring leniency and range restriction may also be related to the scoring system used with these instruments. If the scoring criteria are too vague or difficult to understand or if the rating scale offers too few choices for scoring, then an accurate, fair judgment can be difficult to make, possibly causing raters to grade very highly, so as to not unfairly penalise anyone being assessed.

Scoring leniency and range restriction are significant problems for the utilisation of peer and self-assessment, because if the instruments do not allow for the discrimination of performance, they have little formative or summative value for the students being assessed. In addition, scoring range restriction negatively impacts on the ability of researchers and educators to determine the relationships between peer and self-assessment (specifically the calculation of correlation), which is important for ascertaining their value as part of a triangulated assessment strategy.

One possible solution to this problem is to increase the specificity of the scoring criteria on the instrument, by increasing their number and targeting them at very discrete areas of student performance. This may reduce the problems associated with using only a few, poorly defined criteria and give raters additional cues when assessing various components of student performance. MacAlpine (1999) converted a peer assessment scoring system for an oral presentation from a single letter grade to 4 items that were rated on a 5-point Likert scale and found that students were better able to discriminate performance, although descriptive statistics demonstrating the change were lacking. Another approach was used by Van Duzer and McMartin (2000), who achieved better distribution in scoring on a peer/self-assessment instrument by removing individual items with restricted ranges and high end scores, however, their revised instrument still demonstrated a ceiling effect.

The present study describes an approach where the peer/self-assessment of oral presentations is changed from looking at a few, global components of performance to

multiple, very discrete components of performance and determines the effect of increasing the specificity of these criteria on three aspects of peer and self-assessment: (i) scoring distribution; (ii) correlations between peer and self-assessments; and (iii) impact on written feedback provided by raters on these instruments.

Method

The participants in this study consisted of 98 students in the fifth year (49 in each of two successive classes in 2000 and 2001) of a five-year Master of Physical Therapy (MPT) curriculum at the University of the Sciences in Philadelphia. There was no significant difference in the grade point averages of the two groups when they entered the study. The setting for the study was a course entitled Clinical Simulations, which is a clinical science capstone course in the curriculum. The course was designed in a problem-based learning format, where students were randomly divided into groups of five or six and then randomly assigned to faculty advisors. Each group was then given a complex patient case problem that involved both the learning of new material and the integration of previously learned material. The product of their group work was an oral presentation to the rest of their class, with all the group members participating equally, along with a handout.

At the end of each group presentation, each member of the class audience (the peer assessment group) and each presenter (the self-assessment group) filled out a group presentation assessment. All assessments were of the entire group; there were no assessments of individual performance. Faculty advisors also filled out an assessment form, but their data was not used in this study, because several of the groups were advised by the author, eliminating those assessments as data sources, which meant that those groups could not be triangulated with faculty assessments. The assessment instrument used for the class of 2000 (Appendix A) consisted of five items, each phrased as a specific question. Each item was scored on a 0 (unsatisfactory) to 4 (excellent) scale (for a total of 20 points), and space was left under each item for written feedback by the rater, which was requested but not required. This instrument was based on instruments used in previous studies dealing with the evaluation of oral presentations (Falchikov, 1995; Orsmond *et al.*, 1996; Culbertson *et al.*, 1997; Heylings & Stefani, 1997; MacAlpine, 1999). Initial analysis of the data from these assessments revealed that the scoring tendencies tended to be in a very narrow range, at the high end of the scale, as will be discussed later. The investigator, who was the coordinator for the course, felt that the assessments were not differentiating accurately the varying levels of performance among the presenting groups. For the next iteration of the course (for the class of 2001), a revised instrument was developed (Appendix B) that expanded the scoring criteria from 5 to 25. These items, which were phrased in the form of statements, were directed at more specific aspects of group presentation performance and were based on student outcome expectations the MPT program uses for outcome assessments and accreditation standards. Each item was still scored on a 0–4 scale (for a possible total of 100 points), but it was deemed impractical to ask for written feedback specifically for each item, and so it was requested only at the end of the instrument.

The scores from the two classes were analysed with descriptive statistics and distribution frequencies to examine the scoring tendencies for each year. The relationships between peer and self-assessment were explored using two correlational statistics. Pearson product-moment coefficients (r) are commonly used for determining systematic relationships between groups of assessors (Rezler, 1989; Falchikov, 1995; Orsmond *et*

TABLE 1. Descriptive statistics for peer and self-assessments of student presentations with the initial and revised instruments

Assessor	Instrument		
	Initial[a]	Initial[b]	Revised[c]
Peer			
n	353		430
Mean	19.170	95.850	80.531
SD	1.213	6.065	8.337
Self			
n	49		49
Mean	19.020	95.100	84.449
SD	1.108	5.540	7.472

[a] Scored on a 0–20 scale.
[b] Converted to a 0–100 scale.
[c] Scored on a 0–100 scale.

al., 1996; Heylings & Stefani, 1997; Longhurst & Norton, 1997; Das *et al.*, 1998; Cheng & Warren, 1999; Sullivan *et al.*, 1999). However, if the main area of correlational interest is the level of agreement between groups, and not just association, then an intraclass correlation coefficient (ICC) is more appropriate for analysis (Portney & Watkins, 1993).

The written feedback produced by the peer raters was analysed three ways: (i) the amount of feedback, including the percentage of peer raters who gave feedback and the number of statements on each assessment; (ii) the type of feedback, be it positive (complimentary in nature), negative, (critical in nature) or neutral (comments that could not be discerned as either positive or negative); and (iii) the topic of feedback, concerning either the content of the presentation, the presentation method or general comments.

Results

The descriptive statistics for peer and self-assessment using the initial instrument (class of 2000) and the revised instrument (class of 2001) are found in Table 1. The scores using the initial instrument are presented in their original form, based on a 20-point scale, and also converted to a 100-point scale to allow easier comparison to the data from the revised instrument. The data from the initial instrument clearly shows the tendency for very high mean scores for both peer (19.170/20) and self-assessment (19.020/20); this finding was the stimulus for developing the revised instrument. The revised instrument resulted in lower mean scores for both peer (80.531/100) and self-assessment (84.449/100). The peer and self-assessment scores using the two instruments were significantly different (peer, $t = 29.710$, $P < 0.0001$; self, $t = 8.015$, $P < 0.0001$). The scoring distribution graphs are shown in Figures 1 and 2 (peer assessment) and Figures 3 and 4 (self-assessment). They show that using the revised instrument resulted in a much wider distribution of scores and less tendency for assigning very high scores, indicating that the revised instrument helped the raters in the class of 2001 make more discriminating judgments about the quality of the presentations.

Correlational statistics are shown in Table 2. The Pearson product-moment coefficient between peer and self-assessment was very low using the initial instrument ($r = 0.052$)
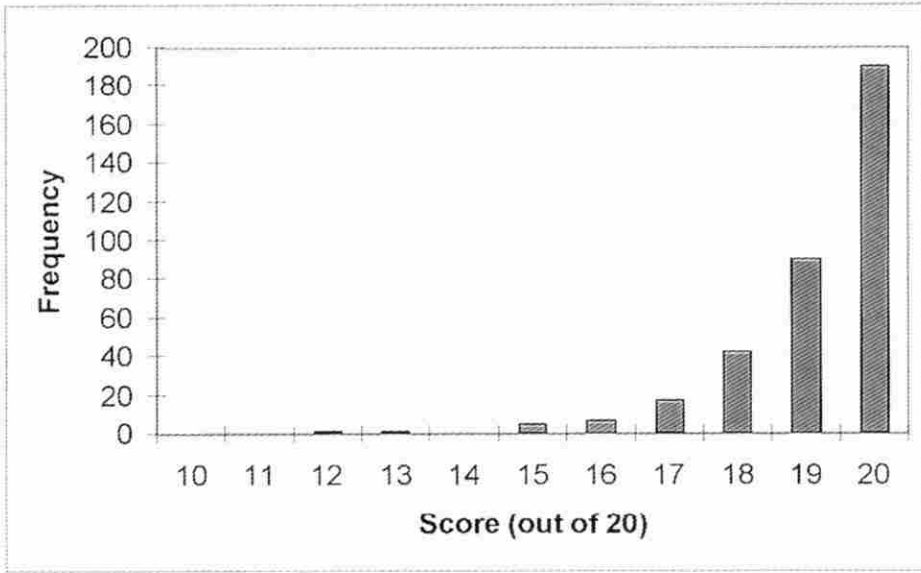
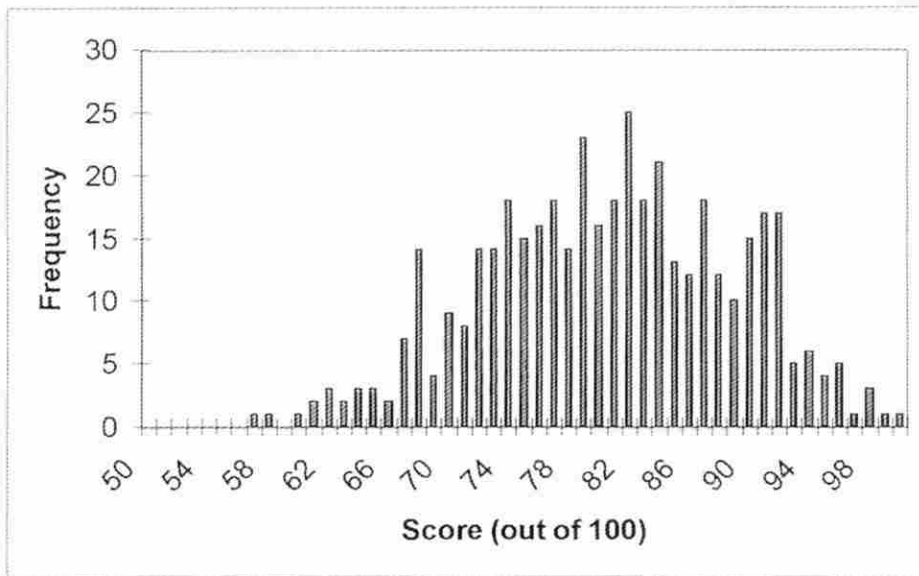FIG. 1. Peer assessment scoring distribution with initial instrument.

FIG. 2. Peer assessment scoring distribution with revised instrument.

but only slightly improved using the revised instrument ($r = 0.215$), indicating a low level of association between peer and self-assessment in both cases. The intraclass correlation coefficient was also very low using the initial instrument (ICC $= 0.175$) but was very high with the revised instrument (ICC $= 0.911$), indicating that the peer and
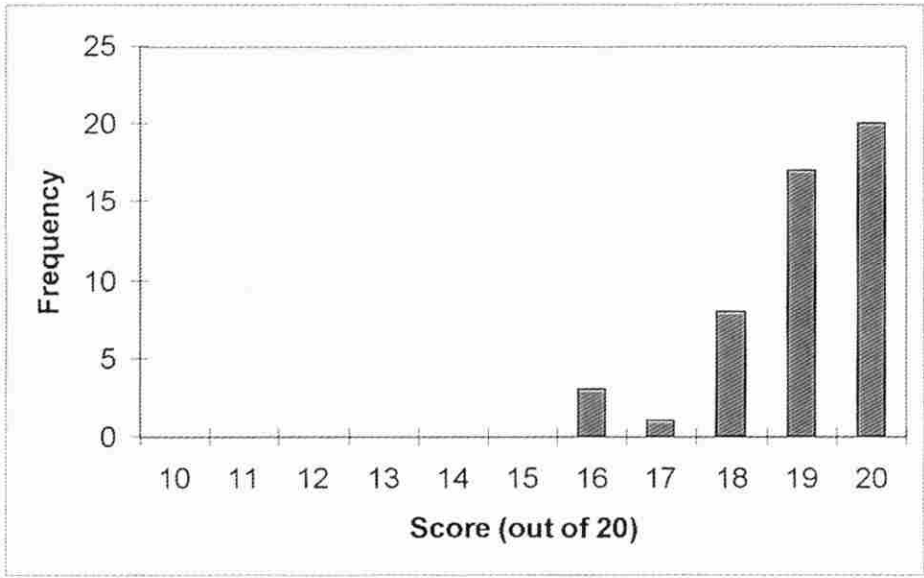
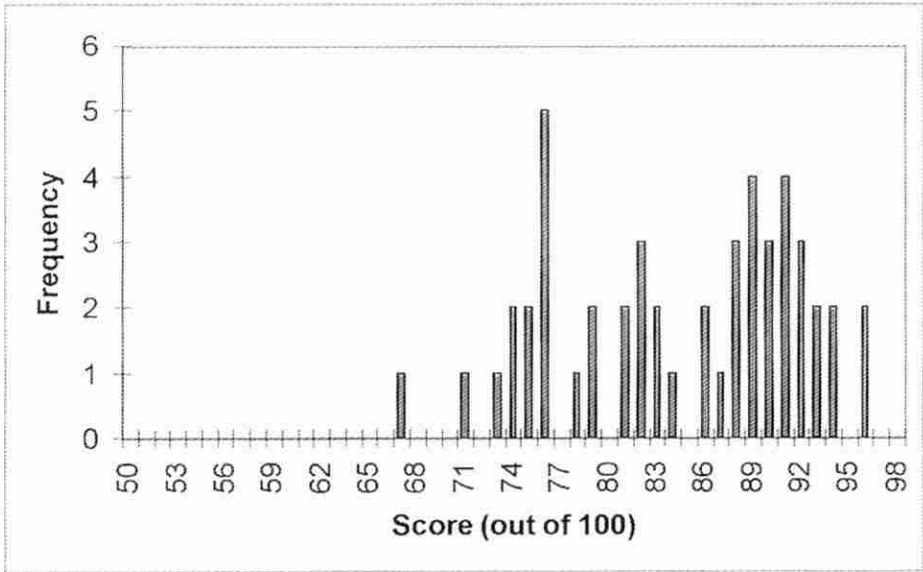FIG. 3. Self-assessment scoring distribution with initial instrument.

FIG. 4. Self-assessment scoring distribution with revised instrument.

self-assessors from the class of 2001 had a very high level of agreement with the revised instrument.

The analysis of the written feedback from peer raters is shown in Table 3. The results indicate that there was a higher percentage of peer raters who gave feedback when the initial instrument was used ($\chi^2 = 263.62$, $P < 0.01$), as well as a higher number of

TABLE 2. Correlational statistics for peer and self-assessment

Type of correlation	Year	
	2000	2001
r	0.052	0.215
ICC $(1,k)$	0.175	0.911

TABLE 3. Comparison of written feedback on peer assessments

	Instrument	
	Initial	Revised
Amount of written feedback		
Percentage of peer assessments with written feedback	95% (336/353)	75% (321/430)
Number of statements per assessment[a]	4.432	3.704
Type of feedback		
Positive comments	81%	74%
Negative comments	12%	21%
Neutral comments	7%	5%
Topic of feedback		
Method of presentation	45%	48%
Content of presentation	45%	40%
General comments	10%	12%

[a] Based on the peer assessments with written feedback.

statements per instrument. The data for the type and topic of feedback were not remarkably different for the two different instruments, with the exception that the peer raters using the revised instrument tended to have a higher percentage of critical comments than those using the initial instrument. It should be noted that data for written feedback for self-assessment is not included, as very few self-assessors wrote feedback.

Discussion

The initial peer and self-assessment instrument (used for the class of 2000) was based on instruments used in previous studies, and the results were similar: there was a clear tendency for raters to assign scores in a very narrow range, concentrated at the high end of the scoring scale. There are several possible causes for this finding.

(a) All of the group presentations were of excellent quality: this is not likely, as the investigator, who attended all of the presentations, found considerable differences in quality among groups.
(b) The raters were not able to determine differences in presentation performance or not willing to be critical of themselves or their peers: this is a possibility and has been raised as a concern by other authors (Falchikov, 1995; Orsmond *et al.*, 1996; Lopez-Real & Chan, 1999; Sullivan *et al.*, 1999; Lejk & Wyvill, 2001).
(c) The instrument did not provide for the discrimination of levels of performance: this

could be due to a lack of specificity of the scoring criteria or the scoring scale not offering enough choices to differentiate performance.

The significant change made to the assessment instrument was to increase the specificity of the scoring criteria, particularly in the assessment of the actual content of the presentations. The revised instrument produced lower mean scores for both peer and self-assessment, as well as a wider distribution of scores. It is the author's opinion that the overall quality of the group presentations for the class of 2001 was similar to the class of 2000, and so the differences in scoring reflect the change in the assessment instrument. The more specific criteria may have allowed the raters to reflect on more aspects of the quality of the performance and, since there were so many items, the peer raters may have felt more comfortable about downgrading certain areas of performance, as it would not have a large impact on the overall score. Other studies have recommended that peer and self-assessment instruments be kept as simple as possible (Oldfield & MacAlpine, 1995; Lopez-Real & Chan, 1999), but the results of this study indicate that a more complex instrument produces better quantitative discrimination of performance, which is necessary if the instrument is to have any validity.

The correlation between peer and self-assessment improved from the initial to the revised instrument, and this can be related directly to the increased variance in scores. The strength of correlations is adversely affected by low variances in sets of scores, as was the case with the data from the initial instrument. Even though the r values were improved with the revised instrument (from 0.052 to 0.215), the association between peer and self-assessment was still weak, a finding seen in other studies (Arthur, 1995; Sullivan *et al.*, 1999). Most correlational studies involving peer and self-assessment are interested in the level of agreement between two sets of raters and because of this the Pearson product-moment may not be the most appropriate statistical tool, since it is possible to have a high level of association with a low level of agreement. Because of that, the intraclass correlation coefficient was also used in this analysis, and the data showed that the increased distribution of scores resulted in a relatively higher level of agreement for the revised instrument.

As was mentioned earlier, the assessments from different raters are frequently used together for the purpose of triangulation. Educators who want to determine the value of a triangulated assessment system must first decide what kind of validity the system should demonstrate. If the purpose is to achieve agreement among the different raters, then the revised instrument used in this study demonstrates the convergence of different raters on a 'single truth' about the quality of the oral presentation. However, triangulation can also serve to uncover the presence of multiple perspectives about the performance being assessed, which do not necessarily have to agree. For this to occur, it is important that the assessment instrument allow the discrimination of varying levels of performance (sensitivity), and the initial instrument in this study did not, while the revised instrument did, even though the peer and self-raters tended to agree with one another.

It should not be taken for granted that different raters should all agree on the quality of a performance. It is conceivable that different groups of assessors may have different expectations of a performance, and this would affect their assessment. To use the context of this study as an example, peer raters may have as their primary expectation that the oral presentation will expose them to new material and judge the presentation accordingly, while self-raters may judge themselves on how well they delivered the information in the presentation. Different expectations would likely diminish the convergent validity

of an assessment, but can strengthen the 'completeness' validity, as long as the differing perspectives are identified.

The results of the analysis of the written feedback showed that the initial instrument, with its directed questions, elicited a greater percentage of peer raters who offered comments and a larger number of comments as well. The revised instrument did not ask for specific responses and this may have been the cause for fewer comments. The larger number of criteria on the revised instrument may have made the raters more critically analytical of the presentations, as they offered a larger percentage of negative feedback than did the raters with the initial instrument.

Implications and Further Study

The results of this study demonstrate that a highly specific assessment instrument, as opposed to a more global instrument, produces better quantitative differentiation of levels of performance at the expense of losing qualitative feedback, which can impact on how the students being assessed view and learn from their assessments. Educators should consider the effects of criteria specificity and written feedback solicitation on rater behaviour when designing these instruments, as they may produce differences in the summative and formative effects of the assessments.

The results of this study indicate several paths for further work, both involving the assessors rather than the assessment. It would be helpful in the further development of a peer/self-assessment instrument to gain a better understanding of what types of feedback students find to have more summative and formative value. It would also be helpful to investigate different assessor groups (faculty, peers and selves) concerning their views on peer and self-assessment, which has not been well-studied (Hanrahan & Isaacs, 2001). Finally, similar studies of this type should be conducted using students at different levels of their education, in different disciplines.

Notes on Contributor

PETER J. MILLER Ph.D., PT is an Associate Professor in the Department of Physical Therapy, University of the Sciences in Philadelphia. He is interested in the use of peer and self-assessment as a formative tool in professional education and also concept mapping as a learning strategy. *Correspondence:* Department of Physical Therapy, University of the Sciences in Philadelphia, 600 South 43rd Street, Philadelphia, PA, USA, 19104. E-mail: p.miller@usip.edu

REFERENCES

ARTHUR, H. (1995) Student self-evaluations: how useful? How valid? *International Journal of Nursing Studies*, 32, pp. 271–276.
BAILEY, D.M. (1997) *Research for the Health Professional*, 2nd edn (Philadelphia, PA, F.A. Davis).
BREITMEYER, B. J., AYRES, L. & KNAFL, K. A. (1993) Triangulation in qualitative research: evaluation of completeness and confirmation purposes, *IMAGE Journal of Nursing Scholarship*, 25, pp. 237–243.
CHENG, W. & WARREN, M. (1999) Peer and teacher assessment of the oral and written tasks of a group project, *Assessment & Evaluation in Higher Education*, 24, pp. 301–314.
CONWAY, R., KEMBER, D., SIVAN, A. & WU, M. (1993) Peer assessment of an individual's contribution to a group project, *Assessment & Evaluation in Higher Education*, 18, pp. 45–56.

CULBERTSON, V. L., KALE, M. & JARVI, E. J. (1997) Problem-based learning: a tutorial model incorporating pharmaceutical diagnosis, *American Journal of Pharmaceutical Education*, 61, pp. 18–26.

DAS, M., MPOFU, D., DUNN, E. & LANPHEAR, J. H. (1998) Self and tutor evaluations in problem-based learning tutorials: is there a relationship?, *Medical Education*, 32, pp. 411–418.

DESMARCHAIS, J. E. & VU, N. V. (1996) Developing and evaluating the student assessment system in the preclinical problem-based curriculum at Sherbrooke, *Academic Medicine*, 71, pp. 274–283.

DENZIN, N. K. (1989) *The Research Act. A theoretical introduction to sociological methods*, 3rd edn (Englewood Cliffs, NJ, Prentice Hall).

DENZIN, N. K. & LINCOLN, Y. S. (1994) *Handbook of Qualitative Research* (Thousand Oaks, CA, Sage).

FALCHIKOV, N. (1995) Peer feedback marking: developing peer assessmentl, *Innovations in Education & Training International*, 32, pp. 175–187.

FALCHIKOV, N. & GOLDFINCH, J. (2000) Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks, *Review of Educational Research*, 70, pp. 287–322.

HANRAHAN, S. J. & ISAACS, G. (2001) Assessing self-and peer assessment: the students' views, *Higher Education Research and Development*, 20, pp. 53–69.

HEYLINGS, D. J. & STEFANI, L. A. (1997) Peer assessment feedback marking in a large medical anatomy class, *Medical Education*, 31, pp. 281–286.

LEJK, M. & WYVILL, M. (2001) Peer assessment of contributions to a group project: a comparison of holistic and category-based approaches, *Assessment & Evaluation in Higher Education*, 26, pp. 61–72.

LI, L. K. (2001) Some refinements on peer assessment of group projects, *Assessment & Evaluation in Higher Education*, 26, pp. 6–18.

LONGHURST, N. & NORTON, L. S. (1997) Self-assessment in coursework essays, *Studies in Educational Evaluation*, 23, pp. 319–330.

LOPEZ-REAL, F. & CHAN, Y. R. (1999) Peer assessment of a group project in a primary mathematics education course, *Assessment & Evaluation in Higher Education*, 24, pp. 67–79.

MACALPINE, J. M. (1999) Improving and encouraging peer assessment of student presentations, *Assessment & Evaluation in Higher Education*, 24, pp. 15–25.

MACPHERSON, K. (1999) The development of critical thinking skills in undergraduate supervisory management units: efficacy of student peer assessment, *Assessment & Evaluation in Higher Education*, 24, pp. 273–284.

MAGIN, D. J. (2001) A novel technique for comparing the reliability of multiple peer assessments with that of single teacher assessments of group process work, *Assessment & Evaluation in Higher Education*, 26, pp. 139–152.

MILLER, P. J. (1999) The agreement of peer assessment and self-assessment of learning processes in problem-based learning, *Journal of Physical Therapy Education*, 13, pp. 26–30.

OLDFIELD, K. A. & MACALPINE, J. M. (1995) Peer and self-assessment at the tertiary level—an experiential report, *Assessment & Evaluation in Higher Education*, 20, pp. 125–132.

ORSMOND, P., MERRY, S. & REILING, K. (1996) The importance of marking criteria in the use of peer assessment, *Assessment & Evaluation in Higher Education*, 21, pp. 239–250.

ORSMOND, P., MERRY, S. & REILING, K. (2000) The use of student derived marking criteria in peer and self-assessment, *Assessment & Evaluation in Higher Education*, 25, pp. 24–38.

POND, K. & UL-HAQ, R. (1997) Learning to assess students using peer review, *Studies in Educational Evaluation*, 23, pp. 331–348.

PORTNEY, L. G. & WATKINS, M. P. (1993) *Foundations of Clinical Research. Applications to practice* (Norwalk, CT, Appleton & Lange).

PURCHASE, H. C. (2000) Learning about interface design through peer assessment, *Assessment & Evaluation in Higher Education*, 25, pp. 341–352.

REZLER, A. G. (1989) Self-assessment in problem-based groups, *Medical Teacher*, 11, pp. 151–156.

SIM, J. & SHARP, K. (1998) A critical appraisal of the role of triangulation in nursing research, *International Journal of Nursing Studies*, 35, pp. 23–31.

SLUIJSMANS, D. M., MOERKERKE, G., MERRIENBOER, J. J. & DOCHY, F. J. (2001) Peer assessment in problem-based learning, *Studies in Educational Evaluation*, 27, pp. 153–173.

SULLIVAN, M. E., HITCHCOCK, M. A. & DUNNINGTON, G. L. (1999) Peer and self-assessment during problem-based tutorials, *American Journal of Surgery*, 177, pp. 266–269.

THOMAS, R. E. (1997) Problem-based learning: measurable outcomes, *Medical Education*, 31, pp. 320–329.

TOPPING, K. (1998) Peer assessment between students in colleges and universities, *Review of Educational Research*, 68, pp. 249–276.

VAN DUZER, E. & MCMARTIN, F. (2000) Methods to improve the validity and sensitivity of a self/peer assessment instrument, *IEEE Transactions on Education*, 43, pp. 153–158.

Appendix A. Initial Peer/Self Assessment Instrument

<u>Oral Presentation Assessment</u>
PT 673–Spring 2000

Group: _____

<u>Grading Criteria:</u>
4–excellent
3–good
2–fair
1–poor
0–unsatisfactory

<u>Clarity:</u> _____
Did the group present their material in a manner that was easy to understand, well-organized, and free from redundancy?
    Comments:

<u>Completeness:</u> _____
Did the group appear to cover all of the relevant topics and issues?
    Comments:

<u>Accuracy:</u> _____
Did the group present material that appeared to be consistent with conventional clinical knowledge, and present support for their views from the literature?
    Comments:

<u>Interaction:</u> _____
Did the group utilize methods to make their presentation interactive and interesting?
    Comments:

<u>Support:</u> _____
Did the group provide supportive materials that were helpful during the presentation, and can serve as a resource in the future?
    Comments:

Appendix B. Revised Peer/Self Assessment Instrument

PT 673 – Spring 2001: Group Presentation Assessment

Please rate this group presentation for the following criteria using this rating scale:

4–excellent
3–good
2 – fair
1–poor
0–unsatisfactory

Group _____

The presentation demonstrated the integration of foundational sciences (e.g., anatomy, kinesiology, neuroscience, physiology, psychology).	0	1	2	3	4
The presentation demonstrated the integration of clinical medicine content pertinent to the case problem.	0	1	2	3	4
The presentation incorporated all aspects of clinical decision making, including the guidelines for decision making, for: examination	0	1	2	3	4
evaluation	0	1	2	3	4
diagnosis	0	1	2	3	4
prognosis	0	1	2	3	4
intervention	0	1	2	3	4
outcomes	0	1	2	3	4
The presentation demonstrated the impact of psycho-social issues on physical therapy management.	0	1	2	3	4
The presentation included a plan for community and/or work re-integration.	0	1	2	3	4
The presentation included patient and therapist safety issues.	0	1	2	3	4
The presentation demonstrated the role of physical therapy within a comprehensive patient management framework, including: communication and consultation with other health care providers	0	1	2	3	4
case management planning with other health care providers	0	1	2	3	4
The presentation demonstrated how wellness, injury prevention, and health promotion can be incorporated into a physical therapy case problem as appropriate.	0	1	2	3	4
The presentation demonstrated how personality differences and cultural diversity may impact the case problem.	0	1	2	3	4
The presentation utilized terminology consistent with that utilized in the *Guide to Physical Therapist Practice*.	0	1	2	3	4
The presentation explained aspects of the case problem that may require delegation to and/or supervision of other health care providers.	0	1	2	3	4
The demonstration incorporated patient/client education as part of the intervention.	0	1	2	3	4
The presentation demonstrated a critical analysis of the research literature regarding the case problem, including: methodology	0	1	2	3	4
results	0	1	2	3	4
implications for practice	0	1	2	3	4
The presentation demonstrated an assessment and intervention plan that is based on research evidence.	0	1	2	3	4
The presentation demonstrated effective interaction with the audience in the discussion of the case problem.	0	1	2	3	4
The handout was effective in supplementing the presentation.	0	1	2	3	4
The group presented their case problem in a professional manner.	0	1	2	3	4

Please add any comments on this presentation on the back of this page. Thank You.