

# Making Students' evaluations of teaching effectiveness effective: The critical issues of validity, bias and utility

Herbert W. Marsh and Lawrence A. Roche  
University of Western Sydney

American Psychologist, November, 1997

The authors address issues of validity and reliability of SETs. They point out that teaching is a complex multidimensional activity, which makes SETs challenging to validate.

Overall, class-average SETs are found to be reliable, as indicated by correlations which range from 0.95 for averages of 50 students to 0.60 for averages of 5 students. Studies attempting to separate teacher effects from course effects have found -0.05 correlation between overall ratings of several instructors teaching the same course, compared to 0.61 for the same instructor in two different courses and 0.72 for the same instructor in two separate offerings of the same course.

In contrast to student evaluators, ratings by colleagues and administrators ratings based on classroom visits have not been found reliable, and do not correlate well with SETs or with each other.