ANALYSIS OF STUDENT PERFORMANCE RATINGS* A RESPONSE

MARGARET F. MUNRO
Associate Professor, Faculty of Nursing
University of Toronto

I have reviewed this study report primarily from the perspective of research methodology and statistical treatment since these are my areas of particular interest.

The topic of performance evaluation addressed by Dr. Angus is one in which educators and supervisors of a practice profession such as nursing are continuously needing to examine and develop. Her study, although limited in scope, has added another dimension to the potential framework for valid evaluation of clinical performance, particularly in the area of differentiation in competencies within and between students.

In reviewing the study report it would have been most helpful to have had more information on the accepted definitions for the five scales of behavior adapted from Tate (1964) as well as how these scales were used to collect the performance data. Since the five scales were given equal weighting in the total instrument, a definition of the performance concepts is essential for the reader to determine his assessment with the outcomes. Similarly, it is important to know the number of faculty persons within which a seventy-five percent agreement on item-scale assignment was considered sufficient to indicate instrument interrater reliability. A standard deviation of .75 on item-scale assignment with scales having a total range of 2.00 could be seen as excessive, particularly if the sample in which it was obtained is small in size. It is not clear if students were evaluated by direct observation, or anecdotal record, on a single occasion, or by a single instructor, or if the data for each evaluation period were a composite of a specified number of observations by more than one person. Since individual performances do vary within a considerable range of "normal for that person", one would hope that the data were cumulative in origin.

It is interesting to note that Angus reports mean scores for the total sample above the midpoint on each performance scale. Such

* Analysis of Student Performance Ratings by Monica Angus appeared in the the Spring/Summer 1980 issue of *Nursing Papers* (Vol. 12, 1/2).

results seem not uncommon in reports of evaluation research and raise two questions in my mind: (1) how likely are instructors to evaluate students at levels below the midpoint on a scale when the descriptors are clearly negative? and, (2) what was the range and the median score obtained on each scale? The concept of developing a cut-off score for acceptable performance in each scaled behavior has exciting possibilities for using quantitative measures to determine student promotion through a program. One possible method of developing such a score would be to determine, descriptively and mathematically, a confidence interval of safety, or minimal competence, below the mean score for each scale once that scale has been sufficiently tested to prove valid and reliable.

Angus has addressed the dilemma of assessment very well. It is indeed difficult to determine whether incompetence is a function of inexperience or is a function of a more basic lack of comprehension. We need to pursue research such as this to become more able to facilitate learning and yet safeguard the standards for professional practice.

The indications of shift in scores between the two years, and within the second year of the diploma program, suggest that there may be a number of intraprogram factors operating: course content in each year, teaching emphasis, clinical practice time and focus, and faculty expectations. Angus has suggested further research using this and other instruments to quantity performance evaluations. I would also suggest adding a dimension of study into the relevance of a school's philosophy to the selected scales for measuring performance.

Dr. Angus has given us an interesting example of the type of clinical performance research which is much needed, not easily accomplished, and challenging to each member of the profession in the outcomes reported. I would look forward to seeing an extension of this project in order to increase the scope of its application.