Potential raters developed rating scales anchored by observed behaviors, scaled for dimension and level. Comparability across regional groups was indicated by scale reliabilities over .97. Attention to recorded observations and separation into dimensions should improve accuracy and facilitate constructive discussion with raters. (The Science Citation Index® (SCI®) and the Social Sciences Citation Index® (SSCI®) indicate that this paper has been cited in over 170 publications since 1963, making it one of the most-cited papers published in this journal.)

Patricia C. Smith
Department of Psychology
Bowling Green State University
Bowling Green, OH 43403

May 12, 1983

“Recommending one treatment of employees over another requires comparison of results measured in different situations on comparable scales. Focusing on observed behavior rather than memory could foster comparability. Moreover, setting individual goals for improved performance requires supervisors to discuss actual behavior with employees. With the National League for Nursing, we developed Behaviorally Anchored Rating Scales (BARS), vertical graphic scales defined and anchored by scaled behavioral examples. Independent groups of head nurses defined dimensions to be rated, recalled examples they had experienced, and rated location of examples anchors along the scales. High scale reliabilities showed comparability across regional groups. The resulting BARS, upon which new observations were to be noted and scaled, proved acceptable to users.

“Lorne Kendall recognized innovative features of the procedure and urged its publication. Other researchers were apparently attracted by potential accuracy of measurement. Some concentrated on minutiae of method while neglecting important note-taking, scaling, and discussion of observations.

“Unfortunately, Lorne did not live to see the happier outcomes we had anticipated: