

FROM THE LITERATURE

In the News

An opinion

Testing

Once in a while an overview appears of the state of the art in a certain area of education. The article by Cees van der Vleuten (1996) on assessment in health sciences education, in the first issue of *Advances in Health Sciences Education* (1996) was one like that, and there are several others.

The Fall 2003 issue of *Teaching and Learning in Medicine* (2003;15(4)) has such an article. It deals with several sources of bias in clinical performance ratings: what else than their competence influences student results on clinical examinations. In 19 pages, provided with 121 references, Williams, Klamen & McGaghie (2003) deal with just about every source of bias that can be distinguished.

Global ratings based on observing convenience samples of clinical performance form the primary basis for appraising the clinical competence of medical students, residents and practicing physicians. In summary, the authors report that raters have a 1- or 2-dimensional conception of clinical performance and that they do not recall details. Good news is reported more quickly and fully than bad news, leading to overly generous performance ratings. On top of that, training has little impact on accuracy and reproducibility of clinical performance ratings. The authors conclude that “clinical performance evaluation systems should assure broad, systematic sampling of clinical situations; keep rating instruments short; encourage immediate feedback for teaching and learning purposes; encourage maintenance of written performance notes to support delayed clinical performance ratings; give raters feedback about their ratings; supplement formal with unobtrusive observation; make promotion decisions via group review; supplement traditional observation with other clinical skills measures (e.g. Objective Structured Clinical Examination); encourage rating of specific performances rather than global ratings; and establish the meaning of ratings in the manner used to set normal limits for clinical diagnostic investigations”.

The definition of clinical competence (the degree to which an individual can use the knowledge, skills and judgment associated with the profession to

perform effectively in the domain of possible encounters defining the scope of professional practice (Kane, 1992)) implies that at least three inferences must be made: judging the quality of the observed performance; drawing general conclusions about the person's ability to perform in a universe of similar situations; and extrapolating from the appraisal context to expected performance in practice situations where the practitioner is not being observed. The following characteristics of clinical competence are addressed: multidimensionality; case specificity; context; quality of clinical ratings; accuracy; (intra and inter)rater agreement; accuracy in estimating competence from a discrete set of observations to performance across a range of tasks and situations; and utility. As cognitive, social and environmental sources of bias in clinical ratings, the authors distinguish: data organization; data recall; frequency of rater observation; focus of attention; rater reporting; leniency and stringency; inflation of ratings; failure rates; rater calibration and training; practitioner (ratee's) characteristics; fixed personal characteristics of practitioners like age and gender; impression management; influence by other evaluators; time pressure and distractions; clinical performance = group effort; and the effects of being observed.

Of course it is not always possible to control all these sources of bias. It is however helpful to be aware of them, and of the systematic influence they may have on the scores that we give our students. If only for that reason, everybody who takes his/her evaluations task seriously should learn from this article!

Jan van Dalen
Associate Editor, *Education for Health*

References

- KANE, M.T. (1992). The assessment of professional competence. *Evaluation and the Health Professions*, 15, 163–182.
- VAN DER VLEUTEN, C.P.M. (1996). The assessment of professional competence: developments, research and practical implications. *Advances in Health Sciences Education*, 1, 41–67.
- WILLIAMS, R.G., KLAMEN, D.A. & MCGAGHIE, W.C. (2003). Cognitive, social and environmental sources of bias in clinical performance ratings. *Teaching and Learning in Medicine*, 15(4), 270–292.