

## What the educators are saying

### Confidence cannot hide weakness in performance

Standardised patients have been used for some time to assess examinees' ability to demonstrate their competencies; now standardised examinees (individuals trained to portray candidates of specific ability levels) are being used to see how variables like confidence affect performance scores. In a comparison of checklist scores after a National Board of Medical Examiners prototype clinical skills examination, standardised examinees trained to be confident were not rated systematically different relative to standardised examinees trained to be insecure. Given the small sample size, the more important conclusion is perhaps that the continued successful use of standardised examinees holds great promise for validity and quality assurance testing in multiple domains.

*Teach Learn Med* 2005;17:9-13

### Simulations may be effective if used wisely and studied carefully

Drawing on the experience of the airline industry, medical education now uses more and more simulation technology. Unfortunately all the excitement over the new technologies has often meant that they are used without adequate consideration of pedagogic principles. A recent systematic review of 109 studies looked at whether medical simulations actually facilitate learning. The overall quality of the research was considered weak, but the best available evidence shows a benefit for simulations when four conditions are met: educational feedback is provided, learners are given the opportunity for repetitive practice, exercises based on the simulation are integrated with the curriculum, and tasks range in difficulty. Amid the usual call for more research there is growing evidence of the effectiveness of carefully implemented simulations.

*Med Teach* 2005;27:10-28

### Evaluation efforts need more samples, not more elaboration

As various licensing agencies worldwide continue efforts to refine the competency requirements of physicians there is a tendency to expand performance rating forms to ensure that each competency is represented. A three item form focused on clinical performance, professional behaviour, and

### Changing the answers on MCQs may improve your score



*Will you improve or harm your score if you change your initial responses in multiple choice question (MCQ) examinations? Examinees tend to believe that changing responses lowers test scores, yet they do change responses, and test scores tend to improve as a result. This paradox may have been explained, at least in part, by a study that presented participants with a series of 100 MCQs and manipulated the response deadline and the difficulty of the question. During a second reading of the questions, participants were allowed to change their initial responses. Performance improved on the second version, but not all errors were corrected. Participants overcame errors that had been driven by the need to respond quickly (for example, typographical errors or misinterpretations), but were not able to correct misconceptions induced by difficult problems.*

*Can J Exp Psychol* 2005;59:28-34

comparisons with other house staff was used to collect assessments of surgical residents (registrars) over the course of three years. Reliability analyses showed that any one item provided just as reliable an indication of performance as did all three items. Increasing the number of items beyond three tends to have little impact, but increasing the number of observations collected

provides much greater benefit. The items chosen may have a steering effect on trainees' learning and behaviour, but evaluators should attempt to collect minimal observations often.

*Surgery* 2005;137:141-7

### Working extended shifts raises the risk of car crashes

Doctors tend to work long hours and extended shifts lasting 24 hours or more. The impact of this practice on medical errors has long been of concern, but little information about the direct link between extended shifts and safety outcomes is available. In monthly reports on their work hours, shifts, and motor vehicle crashes and near misses, 2737 residents (registrars) in the US reported that they spent an average of 70.7 hours in hospital per week and completed 3.9 extended shifts per month, each shift lasting an average of 32 hours. They reported 320 vehicle crashes (82% of which were supported by official documentation). After working an extended shift rather than a non-extended shift, they were twice as likely to be involved in a crash and six times as likely to have a near miss.

*N Engl J Med* 2005;352:125-34

### Only some students need to fill in an evaluation form

Teaching evaluations are useful for both students and faculty. They can provide crucial feedback for teachers, with tenure and promotion often hanging in the balance, and they can be a real voice for students. On the other hand, it is easy to overwhelm students and staff with requests to complete or analyse a large number of evaluations. Instead of surveying the entire class, a sampling approach can be used. During a large clinical practice course, this approach yielded a very high response rate (89%). Generalisability analyses showed that an acceptable level of reliability can be achieved with as few as 10 raters, and the data revealed that the number of items included on the teaching evaluation form have only a negligible impact on the reliability of the ratings.

*Med Educ* 2005;39:171-5

Kevin W Eva assistant professor (evakw@mcmaster.ca) Department of Clinical Epidemiology and Biostatistics, Program for Educational Research and Development, McMaster University, Hamilton, ON, Canada L8N 3Z5