

from the pilot examinations in 2006, after 11 diets in eight separate specialties and with almost 10,000 questions in the bank, there is much to report.¹ The contribution from specialists throughout the UK to this effort has been superb.

Cookson is correct in saying that the pilot examinations were not mapped robustly to the curricula. Furthermore, progressive revision of the specialty curricula during the last two years has presented a moving target for the new examining boards. We have risen to this challenge. From 2009, each SCE blueprint has been mapped to the appropriate curriculum and every usable question related to the relevant curriculum domain. Question-writing groups are giving priority to the remaining gaps.

He criticises the curricula for differentiating between knowledge, skills and attitudes and expresses concern that the SCEs assess only knowledge. Although single best answer questions can evaluate problem-solving skills and clinical judgement, the SCEs were always intended as knowledge-based assessments. They were not designed to test skills or attitudes, which, we agree, are much better evaluated by direct observation and discussion face to face.

Cookson expresses disappointment that the indices of reliability in the pilots were inconsistent. Values of Cronbach's α obtained in examination diets of 200 questions, involving small cohorts with a narrow range of ability, are unlikely to reach 0.9. Indeed, recent research into the use of reliability suggests that the standard error of measurement may be a more appropriate metric.² Nevertheless, it is reassuring that in nine out of 11 SCE diets to date reliability values have exceeded 0.8.

We appreciate the challenge of standard setting for new examinations. For information, the SCEs use the same criterion-referencing process (the Angoff method) used for the MRCP(UK) written examinations in recent years. Although many of those involved in the process had no previous experience, their task was made simpler by taking as a consistent yardstick the knowledge expected of a newly appointed specialist.

JOHN MUCKLOW

*Associate medical director with responsibility for
specialty certificate examinations*

JANE DACRE

*Medical director
MRCP (UK) Central Office*

References

- 1 Joint Committee on Higher Medical Training. *Knowledge-based assessment; pilot project*. London: Joint Committee on Higher Medical Training, 2006.
www.jrcptb.org.uk/SiteCollectionDocuments/KBA%20Project%20Final%20Report.pdf
- 2 Tighe J, McManus IC, Dewhurst NG, Chis L, Mucklow J. The standard error of measurement is a more appropriate measure of quality for postgraduate medical assessments than is reliability: an analysis of MRCP(UK) written examinations, 2002–2008, and Specialty Certificate Examinations. *BioMed Central Medical Education* (accepted for publication).

In response

It is good to read an update and that data continue to be gathered about the performance of the examination. The progress in question writing is impressive but it will be an ongoing task as some questions will perform poorly, some will go out of date and soon many will be remembered by candidates after the examination and maybe passed around.

It is correct that single best answer questions can evaluate problem solving so it seems a pity that a decision was taken to test only factual recall rather than utilisation of knowledge. The maintenance of validity is indeed not helped by the learning outcomes in the various curricula. My problem is not just that only knowledge outcomes are being assessed. They are written so that they mostly require knowledge recall rather than higher order thinking and further many outcomes are not listed under the appropriate heading so that an examination testing only those listed under knowledge will miss important topics. This shows the importance of designing assessment systems along with the outcomes, something previously neglected in the foundation programme.

My anxieties about setting a specialist examination at the lowest level of Miller's triangle remain.

I identified educational impact, cost-effectiveness and acceptability as issues requiring more information and I was hoping this would be forthcoming. It would be a pity I think, if this examination led candidates to acquire most of their learning from books rather than from patients; a survey of their learning strategy would be of considerable interest. There must also now be some robust data on costs. Even if specialists are giving their time freely there is an opportunity cost; if they are writing questions they cannot be doing something else. The publication of robust costings would be a service to all of us who struggle to provide a good examination product for the resources available.

Establishing that standard setting is at the level of the new consultant will certainly be helpful. There must now be considerable information about the consistency of standard setting between the various diets and specialties. Publication of this would help answer the questions raised by the data in the pilot.

JOHN COOKSON

*Professor of medical education and
undergraduate dean
Hull York Medical School*

European Working Time Directive (1)

Editor – McIntyre *et al* provide quantitative evidence for the deleterious effects of the European Working Time Directive (EWTID) on junior doctor well-being in terms of a higher sickness rate (*Clin Med* April 2010 pp 134–7). They correctly speculate that loss of cohesiveness of the traditional medical team is a key causal factor behind this increase in sickness, as well as the resultant loss in actual numbers of junior doctors available at any one time. While these factors are very likely to be implicated, another factor of 'work compression' should be appreciated, ie an increase in task density (the number of tasks per unit time) for every junior doctor due to the absolute decrease in work hours (in addition to the compounding effect of

lack of available colleagues at one time on the ward). Work compression could also reduce job satisfaction and working conditions and further contribute to sickness absence although this is hypothetical and cannot be specifically concluded from this particular study. I would suggest that measures of task density (although difficult to quantify) as an index of work compression should be included in future studies on the effects of EWTD.

Other relevant measures that would have been informative from this study include job sickness rates among more senior medical staff and nursing staff to assess whether the deleterious effects of the EWTD among the well-being of junior medical staff had more far reaching effects on other staff as a consequence. As a final observation, the pre-EWTD sickness rate of having 14% of the total junior medical staff absent on more than one occasion per year seems high (albeit much less than the 38% post-EWTD). Could it be that even a 56-hour working week causes significant work compression (compared to previous longer working weeks) with its resultant adverse effects on well-being, which has been amplified further in the 48-hour week? Strategies and solutions should therefore focus on improvements in task density and other indices of work compression as useful endpoints, as well as the other factors discussed by McIntyre and colleagues in their study.

ANDREW RL MEDFORD

*Locum consultant in respiratory and
general medicine
North Bristol Lung Centre
Southmead Hospital, Bristol*

European Working Time Directive (2)

Editor – We note with interest the findings of McIntyre *et al* (*Clin Med* April 2010 pp 134–7), although we question the conclusions they reached regarding the role of the European Working Time Directive (EWTD) in increases in sickness absence among junior doctors. Not for the first time, the implementation of the EWTD has been linked to negative effects on junior

doctors' fatigue and well-being.¹ However, we believe that the heightened problems should not necessarily be blamed on the EWTD itself. It is our contention that many of the problems have arisen because of the way in which employers responded in their attempts to comply with the EWTD. While the new work schedules may conform to the EWTD's stipulations (eg a minimum daily rest period of 11 hours), they often fail to take into account other parameters which, although not covered by the EWTD, are nevertheless vital considerations in the management of fatigue.

We recently reported the findings of a large-scale survey of junior doctors in which we showed that such additional parameters are important determinants of the likely impact of EWTD-compliant work schedules on junior doctors' fatigue and well-being.² For example, working more frequent on-calls (either at weekends or during the week) was associated with increased psychological strain and work–life interference, while being restricted to only one rest day after working nights was associated with greater fatigue. In support of the EWTD stipulations, we found that working >48 hours per week and short rest inter-shift intervals were both independently associated with increased fatigue. We also demonstrated that working seven consecutive nights was associated with greater accumulated fatigue and greater work–life interference, compared with working just three or four nights.

We would therefore argue that it is difficult to draw any firm conclusions from the study of McIntyre *et al* regarding the cause of increased sickness absence among doctors following the introduction of the EWTD-compliant working time arrangements, without knowing the way in which the new schedule was implemented and what changes in work schedule features were involved.

PHILIP TUCKER

*Senior lecturer in psychology
Swansea University*

SIMON FOLKAR

*Professor emeritus, Swansea University
Professeur invité, Université Paris Descartes*

References

- 1 Murray A, Pounder R, Mather H, Black DC. Junior doctors' shifts and sleep deprivation – The European Working Time Directive may put doctors' and patients' lives at risk. *BMJ* 2005;330:1404.
- 2 Tucker P, Brown M, Dahlgren A *et al*. The impact of junior doctors' worktime arrangements on their fatigue and well-being. *Scand J Work Environ Health* 2010;Online-first:8.

In response to both

In preparing our manuscript we found little formal research into junior doctor welfare. It seems paradoxical that the European Working Time Directive (EWTD) sought to improve this measure without a clear understanding of how this might best be attained. In the face of constant demand, a reduction in trainee hours without increase in numbers permits four scenarios that allow medical care to continue safely. The first is that prior to the reduction in hours there was large inefficiency (>10%). This has not, to our knowledge, been demonstrated. If such inefficiency were not present then, as Medford suggests, either work is being redistributed, omitted or compressed. Significant redistribution of medical work would be needed to cover the loss of nearly one seventh of the medical workforce. While nursing sickness rose during the period covering the introduction of the 48-hour week this finding is confounded by simultaneous alteration to bank staff remuneration which resulted in lesser bank usage. We do not have accurate data on consultant sick leave. Our finding that inpatient mortality and duration of hospital admission were not compromised supports the interpretation that work was not omitted. We may then speculate that work compression and task density must have increased. This effect will be amplified, as we indicated, with higher rates of sickness in a smaller group of trainees. Medford notes that the measurement of work compression and the impact of such compression upon trainees is difficult to assess, however, we found significant psychological stress in two of 15 trainees in a small (non-validated and unpublished) follow-up study. This merits further research.