TOM BENJAMIN, SYDNEY

Hurdles for the psychologically-impaired:

update on the AMA Guidelines and alternative assessments

Assessment of impairment in physical capacity is difficult enough. Widespread adoption of the *"Guides to the Evaluation of Permanent Impairment"* of the American Medical Association is an attempt to inject some standardisation into the task. The *"Mental & Behavioural Disorders"* Chapter (14) adds its own challenges. This article reviews criticisms and proposed alternatives to the AMA Guides and discusses some of the general principles underlying such assessment.

The Paralympics vividly demonstrated that the science of grouping individuals with multiple impairments has progressed to the point that very exciting competitions can result. Paralympians are categorised fairly reliably on their functional capacities such as balance and strength. Yet their other qualities, such as determination, defy such descriptions. Who would have thought wheelchairs could be pushed at comparable speeds to runners over the hilly course of a marathon? It is sobering that the men's 400m one leg amputee record was a mere 10th of a second outside Cathy Freeman's winning time. Yet there are weak links in functioning which place limits on even such capable people. A gold medal is no guarantee of a job, a stable home life, or even personal happiness. The court's daily task is almost the inverse: to weigh up life attainments and make an award on the basis of an inferred overall impairment index.

The assessment method which has overall proved itself in scientific studies is that of simple linear combination. This consists of the sum of a set of variables multiplied by weights as follows: Overall Impairment = $(\text{impairment}_1 \times \text{weight}_1) + (\text{impairment}_2 \times \text{weight}_2) + \ldots (\text{impairment}_n \times \text{weight}_n)$. The mathematical model is merely a way of combining what are basically subjective judgments. It might seem that this only adds mumbo-jumbo without changing the substance of the rating. Research, however, has consistently shown this not to be so.

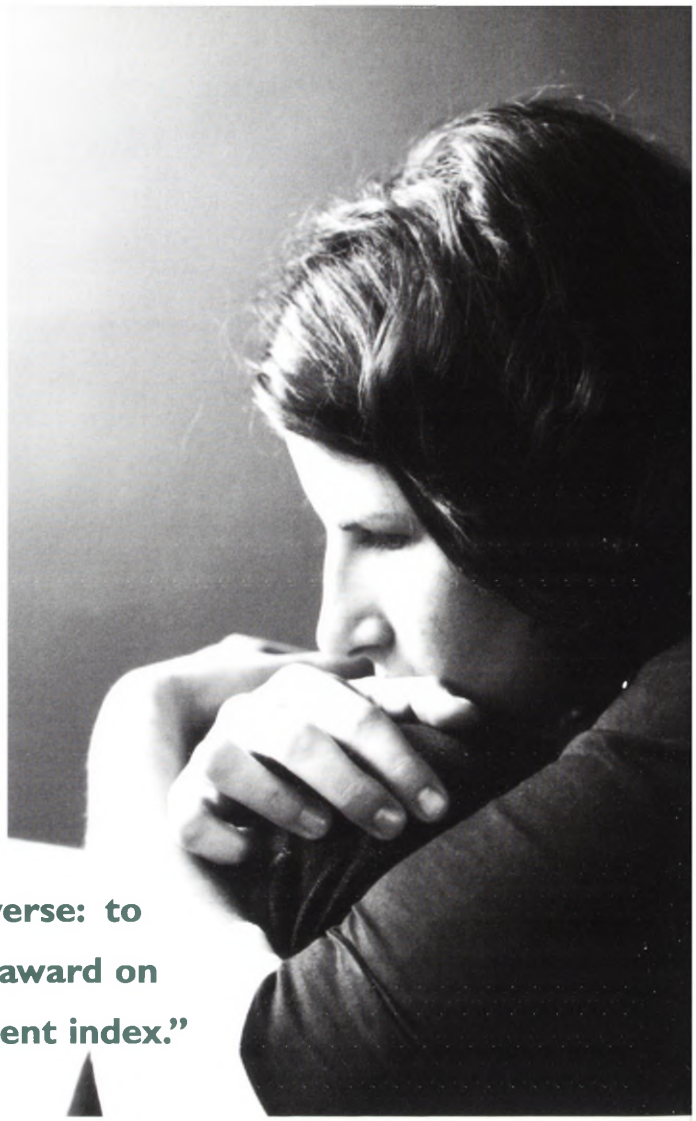
Forty-five years ago, American psychologist Paul Meehl (1954) published a "disturbing" little paper that showed that such linear combinations of ratings could consistently outperform highly paid professionals such as psychologists, doctors and lawyers across all domains of assessment of human ▶

Tom Benjamin is a Forensic Psychologist,
Suite 24, 193 Macquarie Street, Sydney **PHONE** 02 9223 2280
EMAIL tombenjamin@ozemail.com.au **WEBSITE** www.tom.com.au

performance. The findings have been repeated time and again and are so dramatic that clinicians given access to the computer printouts against which they are competing end up adding superfluous interpretations that actually lower their score compared with just accepting the diagnosis in the printout.

Even in the most arcane and highly-paid professions, the answer has been the same - as ace investor Peter Lynch put it in 1993: "fund managers are generally lousy stockpickers and would do better to scrap their computers and throw darts at the business page... Tragically, their residual creativity gets in their way..." In other words, attempts to improve these simple weighting schemes in the long run only lower accuracy.

This lesson, from across the behavioural sciences, is embodied in the AMA Guidelines: a weighted index attempting to capture the totality of a person, much as a bronze, gold, or silver medal summarises the years of dedication, highs, and lows that make up an Olympic performance. The AMA Guidelines for "Mental & Behavioural Disorders" take into account the following broad issues:



"The court's daily task is almost the inverse: to weigh up life attainments and make an award on the basis of an inferred overall impairment index."

- premorbid status
- whether the person has a psychological impairment
- causal link between the accident and the psychological impairment
- whether the impairment is stable
- whether the condition will change as a consequence of time and/or treatment
- impact of impairment daily living functions
- change and level of severity of impact upon daily living

Some criticisms of the approach have to do with the practical difficulties of assessing impact on daily living. This results in low reliability between rates in these areas, affecting the overall score, hence the plaintiff's entitlement. To address these limitations, Australian mental health professionals have proposed alternative models: the Victorian Section of the Australian Psychological Society's Division of Independently Practising Psychologists (1996), a panel of Victorian Psychiatrists (1998), and the APS Working Group on the Measurement of Psychological Impairment (2000) published reviews which largely reflected their separate professional orientations and their recommendations to simplify the approach.

Briefly, the Victorian Psychiatrists' (Epstein et al, 1998) approach was a matrix of six criteria, each rated on a 5 point scale, with whole person impairment calculated as the median rating across the 6 individual criteria, which are: intelligence, thinking, perception, judgment, mood and behaviour. It then incorporates a formal Mental State Examination as the "prime method of evaluating psychiatric impairment" and a formal diagnosis, in line with the Diagnostic and Statistical Manual, 1994 (DSM-IV). The guidelines then include a number of contextual issues that are supposed to also be considered, regarding education, financial, social and family circumstances, motivation to improve, and treatment progress. The overall impairment index was also then to be compared with the Global Assessment of Function scale (GAF), a uni-dimensional scale which forms Axis V of the DSM-IV. Essentially, the psychiatrists' approach was an operational definition of the familiar psychiatric clinical interview.

The psychologists (APS, 2000) did not feel that this approach improved on the original AMA guidelines. Perhaps surprisingly for a professional lobby group, they did not recommend an approach based on their own stock in trade, extensive psychometric tests, but opted for one of the psychiatrists' tools - the Global Assessment of Function. In their view, an index of change in GAF scores could address the fundamental question posed in the impairment assessment. The index recommended was the common formula for percentage change:

"% Functional Impairment Loss = [GAF pre-morbid - GAF Post accident] x 100 / GAF pre-morbid"

This model assumes that the GAF encompasses an "overall measure of a person's level of psychological functioning", incorporating weak links in key areas, and that the aim of the assessment at the end of the day is to gauge whether the person has dropped to a lower category of life functioning overall.

The APS group cited the GAF's known reliability, validity, and widespread use as reasons for their endorsement.

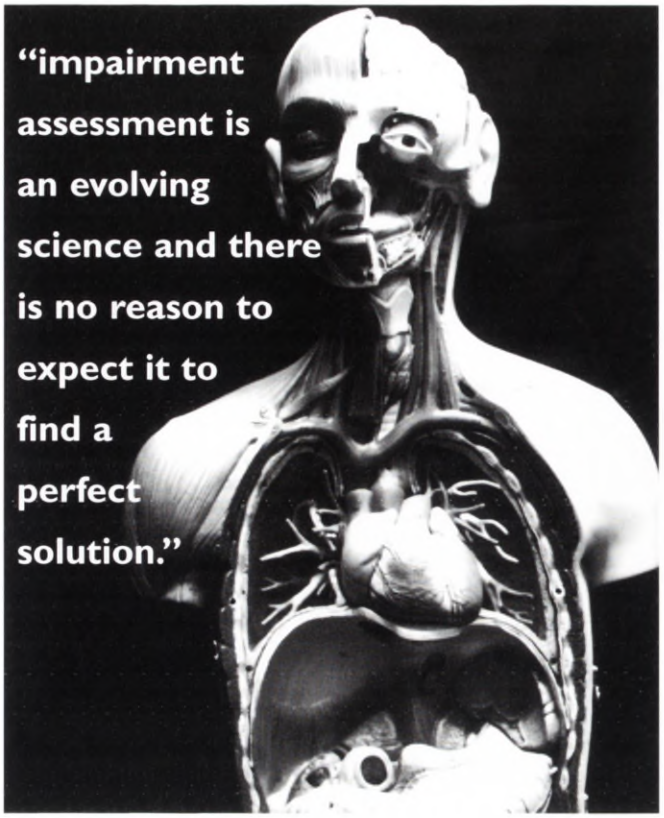
How these methods stack up in the assessment competition has largely to do with their sensitivity, selectivity, and specificity: ie. how many people get put into different classes using these different methods. To give this perspective, I contrast these American and Australian approaches with one of British origin - the "Health of the Nation Outcome Scales" (Wing et al, 1996), developed to "measure the range of physical, personal, and social problems associated with mental illness". Like the scales discussed previously, the HoNOS scale employs a matrix of scores on a 0-4 scale across dimensions of behavioural problems, impairment, symptomatic problems, and social problems. What distinguishes it from these other applications is that its target population of identified "mentally ill" persons allows for somewhat extreme behaviour to fall into the class of "minor problem requiring no action".

For example, a person with the following rated behaviours would thus be rated "minor problem requiring no action":

> Irritability, quarrels, restlessness, Fleeting thoughts about ending it all, Drink or drug over-indulgence, Problems with memory or understanding, Health problems, Odd or eccentric beliefs not in keeping with cultural norms, Gloomy, Minor mental/behavioural problems, Minor relationship problems, Untidy, disorganized, Transient accommodation, Late unemployment cheques.

Imagine the hue and cry if such a person were awarded a child custody or an employment re-instatement, there were then some sort of tragedy, and this psychological report fell into the hands of the press. To inject a bit of reality into such a rating system in the courtroom would surely tempt some judicial luminary to pose a simpler test such as "should this person be allowed to drive the Clapham Omnibus?" In modern times, even this might be deemed a mere armchair judgment. It invites a reply to the judge: "Your Honour, would you hire this person as your Associate or member of chambers?" or more personally, "Your Worship, would you let this person marry your son or daughter?"

Most of us have learned this lesson the hard way. The technical term for it is "ecological validity" - the ability of tests to predict real-life performance. I'm glad I don't wear hats for the number of times I would have had to eat them when patients performed either far better or worse than our expectations. For example, I can recall a client in rehabilitation who passed all of our cognitive and physical tests with ease. Our team conducted detailed work visits, had the full cooperation of his employers, and an enthusiastic, mildly-disabled client. He returned to work but was sacked within the week. Why? He couldn't keep his hands off female colleagues. This hadn't been something we routinely tested and he'd given no inkling of this around hospital female staff. ▶



In summary, impairment assessment is an evolving science and there is no reason to expect it to find a perfect solution to what remains essentially a socio-politico-economic problem rather than a medical one. Settling on a single standard such as the AMA Guidelines for convenience has all the charm and inspiration of the adoption of Microsoft *Windows* as the de-facto universal PC operating system. Many books have been written about *Windows* and the way new "bugs" are introduced into the code of such a complex programme for each batch of old bugs that is eliminated. Microsoft programmers have to resist "feature creep" - the temptation to put new gimmicks and icons in the next versions. So it must be, albeit on a smaller scale, with the AMA Guidelines. Shortcomings in reliability and anomalous assessments will invariably lead to further cries for reform and more tinkering at the edges.

It is my hope that this article will introduce legal professionals to the underlying scientific issues such that a mere technical solution ie. adopting a particular scale such as the AMA Guidelines or GAF - is not seen as a permanent scientific solution. Science requires competition to achieve "paradigm shifts" (Kuhn, 1970). The paradigm here is that of actuarial combination in human judgment, which has been extensively studied. There are many more competing scales available than could be discussed here: the Life Skills Profile, Rand SF-36 and the General Health Questionnaire, just to name a few. The paradigm has demonstrated time and again that, whichever of these are used, the result is likely to be a vast improvement on unstructured professional judgment. The temptations that have to be resisted are those of adding so-called improvements which end up watering down their effectiveness and, more importantly, those of abrogating legal, political, and social responsibilities by locking them into technocratic solutions. PL

References:

American Medical Association (1995), Chapter 14, Mental and Behaviour Disorders, *Guide to the Evaluation of Permanent Impairment,* Chicago, American Medical Association.

American Psychiatric Association (1994), *Diagnostic and statistical manual of mental disorders*, (Fourth (DSM-IV) ed.), Washington DC, American Psychiatric Association.

Australian Psychological Society (2000), *Impairment Measurement Of Psychological Impairment In Matters Of Civil Litigation*, Carlton South, Victoria, APS Working Group on the Measurement of Psychological Impairment.

Epstein, M., Mendelson, G., & Strauss, N. (1998), Guides to the Rating of Permanent Psychiatric Impairment, *Victorian Government Gazette*, *No. S87* (28 August).

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago and London.

Lynch, P. (1993), *Beating the Street*, New York, Simon & Schuster.

Meehl, P. E. (1954), *Clinical Versus Statistical Prediction: a theoretical analysis and review of the literature,* Minneapolis, University of Minnesota Press.

Wing, J., Curtis, R., & Beevor, A. (1996). *HoNOS: Health of the Nation Outcome Scales, Trainers Guide*, London, Royal College of Psychiatrists.