Standardisation versus Innovation in Psychological Testing

Brian Tiplady

For presentation at IEE meeting 18th May 1994

Many studies on the effects of drugs on psychological performance have been carried out in the past few decades. The number of different tests of performance used have been almost as great. They range from measurements of reaction times to the speed of putting tops on ball-point pens; from the rate of crossing out the letter "e" in a page of letters to assessments of standing steadiness; from memory of lists of words to the ability to carry out mental arithmetic (Hindmarch, 1980).

Does such variety simply represent the diversity in the skills that need to be assessed? Or is it rather due to a lack of coherence and focus in research strategies? The latter view is supported by formulations such as those by Steinberg (1959) and Wesnes et al. (1987) who see performance skills as represented by only a few domains. Thus Wesnes and co-authors classify such skills under four main headings, viz: Attention, Memory, Psychomotor; and Cognitive. These domains may of course be subdivided: Focussed and divided attention; short and long-term memory; perceptual speed and speed of response-formation; and so on. But such an approach does suggest that rather few tests will capture a great deal of what needs to be measured

This idea is borne out when we look at some of the actual tests in use. Thus choice reaction time is measured in a wide variety of ways, with different numbers of buttons to press; buttons arranged in circles, squares and semi-circles; buttons operated by one finger for all buttons or by one finger per button and so on. It is difficult to believe that all these variants measure different and important aspects of psychomotor performance!

Where researchers have used standardised measures, it has greatly enhanced the usefulness of the data obtained. An example of this is one particular version of choice reaction time, part of the Leeds Psychomotor Tester. Hindmarch has presented data from a wide variety of drugs on this measure, together with results from the other test on this device, Critical Flicker Fusion (see e.g. Hindmarch et al. 1991).

The fact that all these drugs have been assessed using the same methodology is a great advantage, but a limitation of much of this work is the use of single doses in many of the studies. However the value of the standardisation of methodology such studies is that comparison can be made not just with other studies from the same laboratories, but with work by other groups. The author and co-workers have used the same methods in a number of dose-response studies that lend themselves to direct comparison with the studies cited above.

The first study to consider assessed the effects of nitrous oxide in doses of 5-40% by inhalation (Fagan et al., 1994). This showed a curvilinear dose response, with Choice Reaction Latency showing significant effects at 10% nitrous oxide, while Motor Time showed effects only at 40%. Both measures showed accelerating impairment with concentration. A subsequent study (Armstrong et al., in preparation) investigated the lower end of the curve, looking at concentrations of inhaled nitrous oxide from 3-15%. We applied curve fitting to the combined data to obtain a composite dose-response to which other drug studies can be compared. For example Hindmarch et al, (1991) have studied the performance effects of ethanol at doses up to 1 g/kg. Using the suggestion of Fagan (1991) that 1 mg/kg ethanol is roughly equivalent in its effects to 20% nitrous oxide, we have shown that the curves for ethanol and nitrous oxide show a good match for the two choice reaction time measures. Other drugs may readily be matched to such profiles. Only on the basis of such dose-response comparisons is it possible to determine whether different drugs have distinct profiles of effects on performance (Fagan, 1991).

The use of standardised systems is therefore valuable, and tests of this sort should be included in more studies of drug action (see e.g. Echeverria et al., 1989; Englund et al., 1987). However the use of only such standard tests would lead to a different set of problems. For much has happened in

cognitive psychology since the Leeds Psychomotor Tester was devised. And some of the older tests that have not become standard still deserve a look.

One area where much work has been done is attention. Work in neuropsychology has indicated that functions such as disengagement of attention, orientation to a stimulus, and focusing attention on a stimulus to avoid distractions are distinct, and involve different brain areas (Posner and Petersen, 1990). The time taken for the processes involved in shifting attention from one source of information to another is of great practical importance, e.g. in driving, where a sudden shift in attention may be necessary in an emergency. Measuring the speed of such a shift is not as straightforward as choice reaction time.

One approach to this has been described by Sperling and Reeves (1980). Subjects fixate on a central point on a monitor screen. To one side of this appears a stream of letters, on the other a stream of numbers. Without moving their eyes, subjects attend to the letter stream until a particular target occurs, then attempt to report the first number appearing in the number stream. By taking a substantial number of trials, and varying the time intervals between letters and numbers, a mean time necessary for switching attention from one channel to the other can be obtained. While this test has been used in investigations of attention, it has not apparently been used in psychopharmacology.

This test is of interest because it assesses a process of theoretical and practical interest not normally investigated in drug studies. Another reason for innovation is to improve the sensitivity of test measures. This is part of the rationale for the introduction of tests of divided attention (see e.g. Moskowitz and Sharma, 1974). The Shift of Attention Task may readily be programmed on a standard computer systems (IBM compatible etc). It may also be of interest to introduce new technology into the area. An example is pen-based computing, which has great potential in automating pencil-and paper tests

(Tiplady, 1994). There are also test systems involving CD-ROM interactive graphics, and those which combine electrophysiological measures such as EEG with performance tests.

Thus the relationship between standardisation and innovation is one of balance. But innovation should only be for a good reason - there are many examples in the literature of test variants which add little except non-standardness! And all too few cases where tests altered for a good reason have been compared to the original. Please use at least some of the old faithfuls!

References

Echeverria D, Fine L, Langolf G, Schork A, Sampaio C (1989) Acute neurobehavioural effects of toluene. British Journal Ind.Medicine 46:483-495

Englund CE, Reeves DL, Shingledecker CA, Thorne DE, Wilson KP, Hegge FW (1987) Unified tri-service cognitive performance assessment battery. Naval Health Research Center Report No. 87-10, San Diego, CA, USA

Fagan D (1989) Pharmacokinetics and pharmacodynamics of centrally acting depressants in man. PhD thesis, University of Edinburgh.

Fagan D, Tiplady B, Scott DB (1994) Effects of ethanol on psychomotor performance under steady-state conditions J Psychopharmacol (in press)

Hindmarch I, Kerr JS, Sherwood (1991) The Effects of Alcohol and Other Drugs on Psychomotor Performance and Cognitive Function Alcohol & Alcoholism 26:71-79

Moskowitz H, Sharma S (1974) Effects of alcohol on peripheral vision as a function of attention. Human Factors 16:174-180

Posner MI, Petersen SE (1990) The attention system of the human brain. Ann Rev Neurosci 13:25-42

Sperling G, Reeves A (1980) Measuring the reaction time of a shift of visual attention Attention Perform 8: 347- 360

Steinberg H (1959) Effects of drugs on performance and incentives. In: Lawrence DR(Ed), Quantitative Methods in Human Pharmacology & Therapeutics. London: Pergamon, 1959 : pp 76-87

Tiplady B (1994) The use of personal digital assistants in performance testing in pharmacology. Demonstration, British Pharmacological Society, London, January 1994

Wesnes K, Simpson P, Christmas L (1987) The assessment of human information-processing abilities in psychopharmacology. In: Hindmarch I, Stonier PD (Eds) Human Psychopharmacology. Methods and Measures Vol I. Chichester: Wiley : pp 79-91