Computer Based Testing

Information technology is having a revolutionary impact on psychometrics. The first stage was the application of the number-crunching power of mainframe computers to the necessary statistical calculations of item analysis, reliability estimation and more general experimental design. More recently the computer has been able to take over the role of test administration, and here, because the machine is able to make decisions so quickly, a new stage of development has been reached, where results obtained at the beginning of a testing session can be analysed in time to modify the progress of the session. Artificial intelligence using both expert systems and neural technology is also applied to psychometrics, amid other exciting prospects in terms of their ability to deal with non-linear data and to recognize complex patterns.

The QNET Software Engineering Research Laboratory is currently engaged in extensive research in the potential for computers to model decision processes at a linguistic as well as an information-processing level; which is one of the most productive and controversial areas of development not just in psychometrics but in social decision making in general as well.

Computerized administration of tests has been used for some time, with several advantages over paper-and-pencil testing. Individual test administration in particular is a time-consuming yet fairly automatic process for many tests, so that computerization offers potential savings in professional time, Klinger *et al* (1996) have shown that when used in clinical settings there is also an increased acceptance by patients. It seems that often people are happier to answer questions to a computer than they are to their fellow human beings!

However, the mere placing of a test on a computer for easier administration is only the first step in the utilization of computers for test administration. Because of their speed and infallible memory, computer programs can easily be written to allow the computer to adapt its questioning to the initial responses of the respondent. In a paper-and-pencil test we have no way of knowing how a respondent would respond, so that all eventualities could be covered. On the other hand, with a computerized adaptive test, any questions that turn out to be irrelevant, and which could be altered to retain the effectiveness of the test, is effectively carried out by the system. It is also possible to tailor the difficulty level of a test to the ability level of an applicant. If the respondent encounters difficulty, the test can be made progressively easier, and if the respondent finds the test too easy it can be made more difficult. This procedure tends to increase the motivation of the respondent in either case. It is further possible to target special areas of the blueprint and allow the program to select questions that focus on a particular area.

Thus if we are interested in whether a person scores above a particular cutoff in a test it is possible to use a Bayesian approach which always chooses as the next item one which would maximize information towards making a decision, and to stop as soon as a specified probability level of acceptance or rejection is reached. Adaptive testing seems such an obvious way forward that it is surprising it has not become more widely used. There are several studies that have shown it to be more successful than classical techniques for both criterion-referenced (Haladyna and Roid 1983) and norm-referenced tests (Hambleton and Swaminathan 1985). Reasons for the technique in not becoming universally popular could include the lack of software systems, the sophistication needed to interpret the item response theory models involved, and the lack of faith of educational and psychological professionals in models they cannot understand. The QNET Software Engineering Research Laboratory is proud to be in the forefront in this sphere of research and development of computer based interactive tests.