Standing up to Falling Down – Using the Familiar to Catch the Unusual

Dr Russell Beale
Sam Davies
School of Computer Science
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK
R.Beale@cs.bham.ac.uk

Abstract

Detecting impairment is particularly difficult in people, and is especially important if they are in charge of a vehicle. We describe how we have augmented the familiar and unchallenging medium of pen and paper by using a digital pen coupled with established paper tests in order to develop a screening device for driver impairment which may be used at the roadside. This form of impairment testing does not isolate or intimidate any member of the general public as a computerized test may do, and proves to be highly acceptable and accurate. Results presented show that impairment can be detected, but that the current tests are not discriminatory enough.

Keywords

Impairment, user testing, novel interaction, pervasive computing

INTRODUCTION

Pen and paper is nearly 4,000 years old and after years of evolution are still one of the best ways to collect, store and distribute information. Until recently the only way to digitally present what you have written with a pen has been to scan the image in or use hardware that directly writes to the screen. Anoto, a Swedish company, have developed a technology that allows you to write onto normal paper and digitally stores the x-y co-ordinates traversed. The digital information can then be processed either immediately by sending the information to a server, or subsequently via direct connection to a PC. The initial use of the technology was aimed at the teenage picture texting market, but as the potential for other applications has become clear a wide range of solution are being developed.

This technology has been used in a project in response to a police requirement for technology to assist officers at the roadside in determining whether an individual is impaired though drugs. The prevalence of drugs in modern society has increased dramatically in recent years and consequently a growing proportion of the population is now likely to drive under the influence. It is well documented that alcohol and drugs can act alone and in combination to impair cognitive functions, especially those related to driving.

The driving task requires the human brain to operate on three distinct control levels (Sanders, 1986). These are defined as the strategic level, the manoeuvring level and the control level. These three levels are interconnected and so on any one journey the brain will pass a great deal of information between the three levels. As a driver becomes more experienced, there will be a general shift of actions from the strategic level, down to the control level. A person's ability to manage cerebral resources and the transfer of information between the three control levels is affected by alcohol and drugs. The manner in which these chemicals influence a person's performance in the driving task is incredibly varied. Alcohol alone produces a diverse range of effects such as disrupted balance, impaired vision, slowed reaction time, degraded learning ability and memory (White et al., 2000).

The increase in drug driving is in contrast to the general decrease in 'drink driving' that is, driving with a Blood Alcohol Concentration (BAC) greater than 80mg/100ml or 35µg/100ml breath alcohol concentration. UK

Government legislation and Road Traffic Law is very clear on 'drink driving' and is enforced under Section 5 of the road traffic act 1998. It states;

"If a person drives, attempts to drive or is in charge of a mechanically propelled vehicle on a road or other place after consuming so much alcohol that the proportion of it in his breath, blood or urine exceeds the prescribed limit he is guilty of an offence"

UK Road Traffic Act 1988, amended 1991

Clear legislation, coupled with the deployment of approximately 10,000 alcohol breathalysers throughout the country, means that the UK probably has the most effective technology possible to combat 'drink driving'. As a result of this, the emphasis is now shifting towards 'drug driving' and 'fatigued drivers'. The problem faced by the police is there is no clear law which makes it illegal to drive a car while under the influence of any level of drugs, illicit or otherwise, in the bloodstream. It is illegal to posse illicit drugs and it is illegal to take illicit drugs, but as soon as the drug is within your blood stream you cannot be charged. Also, there is no equivalent to the breathalyser for drugs; any test must be carried out in laboratory conditions with expensive equipment operated by qualified personnel.

If a police officer stops a driver whom they suspect of being under the influence of drugs, the officer then has the option of carrying out Field Impairment Tests (FIT). This suite of tests involves a pupillary examination, asking the suspect to walk in a straight line, touch the end of their noise with their finger, a one leg stand and estimation of 30 seconds. A full description of FIT can be found in O'Keefe (2001). The FIT battery is only used as a screening technique to allow the officer to assess the individual's level of impairment. If the driver is still suspected of driving under the influence then they must be taken to the local Police station to be subjected to tests by a Police surgeon. This process generally takes over three hours, by which time any affect that the drugs had, may have worn off.

Section 4 of the UK Road Traffic Act is used to prosecute people who drive under the influence of alcohol and drugs;

"Any person who, when driving, attempting to drive or in charge of a mechanically propelled vehicle on a road or other public place is unfit to drive though drink or drugs is guilty of an offence."

UK Road Traffic Act 1988, amended 1991

To prosecute an individual under section 4, it must be proven that they were "unfit to drive through drink or drugs". This raises a number of important questions, principally: What does unfit to drive mean and how do you show that someone was unfit to drive?

Although the number of people killed on the UK roads has decreased annually by almost 40% since the 1970's, on average 10 people are still killed everyday in Road Traffic Collisions (RTCs). A report published by the Transport Research Laboratory (TRL), showed that 18% of people fatally injured in RTCs during 1998 had measurable quantities of illicit drugs in their systems (Tunbridge et al 1998). This figure represented a six-fold increase from a similar study carried out by Everest et al. (1989). It is likely that the proportion of road deaths that are drug related will have increased further based on social trends and an increasing number of police reports of drugged driving.

To date numerous computer-based test batteries have been developed to assess an individual's level of impairment (Hope, A et al (1998), Clark A et al (2000) and Cameron, E (2001)). The majority of impairment tests call upon a user to display general psychomotor skills, including things such as reaction times, divided attention and visual tracking. To provide a portable testing unit, platforms such as Pocket PC's or Newton Message Pad have been employed.

Computerised test packages are attractive because they offer the potential for standardised, sensitive and quantitative measures of psychomotor performance. A major problem with such tests, however, is that not only are an individual's psychomotor skills being tested, but so too is their ability to interact with a machine. Individuals with computer experience may have an advantage over users who have had little interaction with this form of technology. Earlier informal work by one of the authors showed a noticeable difference in the performance of computer-based tasks between males and females, probably due to the increased familiarity of men with computer games and similar tests. Another important consideration is the screen size that will affect how well a subject is able to perform the tests. Many Pocket PCs may provide convenient platforms for the installation of impairment tests but their small screen sizes may pose difficulties for some users and may reduce the overall sensitivity of certain types of test.

OBJECTIVES

The primary aim for this project was to develop a way of determining impairment, without using technology that might discriminate against certain user groups. Pen and paper is recognised as a natural form of interaction that is familiar to users, where prior exposure to computers is not a benefit. Pen and paper tests have been used in drugs trials to determine an individual's level of impairment for many years. Tests such as the Gibson Spiral Maze, Digit Symbol Substitution and Trail-Making have been extensively used to investigate the psychomotor and cognitive effects of drugs and other conditions affecting CNS function.

One problem with using these methods in the past has been variability in scoring the tests. Although there are sets of established rules, the way these are interpreted differs from user to user. It is also a time consuming and tedious job to hand mark large numbers of tests, which is not a practical activity for police officers.

The development of a digital pen and paper system (Anoto®) that records exactly what the user has written may help solve the traditional problems associated with pen and paper testing. The digital pen writes on ordinary paper printed with a unique pattern that is almost invisible to the naked eye. Figure 1 shows an enlarged view of the paper grid. A tiny camera inside the pen registers the movement across the grid surface on the paper and stores it as series of map coordinates at a rate of 100/second. The unique displacement of dots makes it possible for the pen to determine the position within the page.

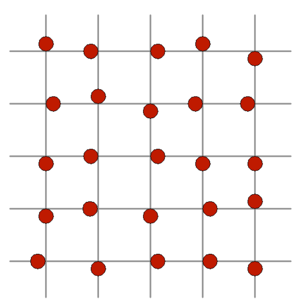

Figure 1: Close Up of
Anoto Pattern

The pen uses a camera and a tiny LED to take snap shots between 50 – 100 times per second of the pattern printed on the page. To be able to calculate the pens position the camera reads a 7mm square area around the tip of the pen. This allows a user to orientate the pen at any angle and still record the same positional data, shown in Figure 2.

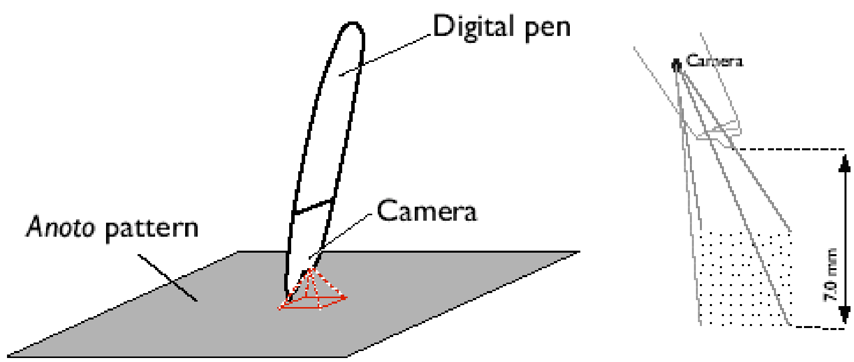

Figure 2: Digital Pen Camera Range

The pen processes the snapshots by calculating the absolute position on the page and saving them to memory to be later transferred as pen strokes to an application.

The pen (Figure 3) feels and works exactly the same as a normal pen, albeit a little larger. It connects with a peripheral device via either a docking station or Bluetooth. Once the series of coordinates has been transmitted it can be reconstructed into an exact copy of the handwriting for display. The accuracy is estimated at 0.3mm (i.e. the distance between adjacent grid points) and the pen can store up to 40 A4 pages of continuous writing. In addition the pen records time and pressure information at each data point.

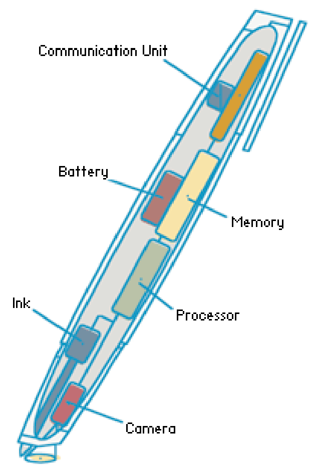

Figure 3: Logitech digital pen

With the use of the Anoto Software Development Kit (SDK) and a Logitech digital pen we created a demonstration package to evaluate the potential for the technology to be used operationally at the roadside by the police.

USER TESTING

Two tests were selected as the most effect for demonstrating the pen's potential:

Tiplady Tracking Test (Figure 4): This test of psychomotor speed and accuracy is similar in principle to the Gibson spiral maze (Gibson H. B. (1978)). The user follows a zigzag track with the pen as quickly as possible while avoiding the edge of the track and obstacles. Total time is recorded, and an error score calculated on the basis of the number of times the pen touches or penetrates an obstacle or the edge of the track. The test is scaled to fit onto one sheet of A4 paper.

Trail-Making Test: The Trail-Making Test (TMT, see Figure 5) is one of the most popular and most widely used tests in neuropsychological research and assessment (Reitan, R. M. (1958)), although it has not been so widely used in drug research. Volunteers are asked to join the numbered or lettered shapes up in the correct consecutive order as quickly as possible. This test is significantly different from the tracking test, having a greater cognitive component, and it further demonstrates the potential of the pen system.

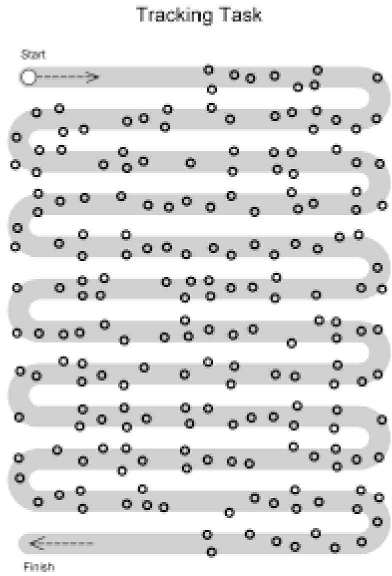

Figure 4: Tiplady Tracking Test

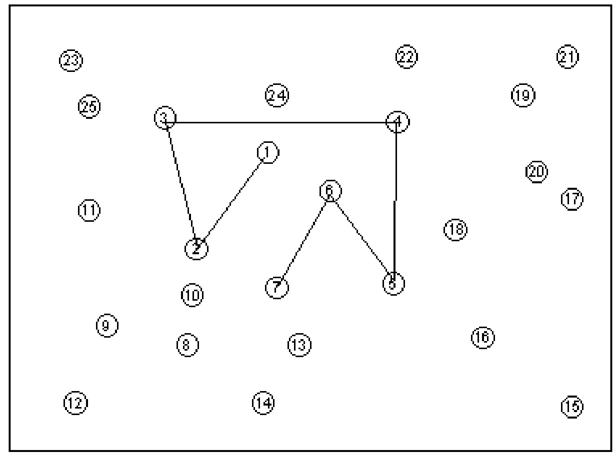

Figure 5: Trail-Making Test

Consideration was also made about including a digit substitution test, but it was felt that character recognition software is not yet at the level to offer any distinct advantage.

Methodology

The purpose of digitizing the majority of impairment tests is to help improve accuracy. Whereas with a traditional pen and paper test the marking scheme is likely to alter from user to user, this solution will allow the tester to be confident every individual has been evaluated in the same way.

The software was designed to evaluate the Anoto functionality and assess its suitability for paper and pen impairment tests. To date the Zigzag Tracking and Trails Tests have been programmed to work with the Logitech digital pen. The software is designed to allow for data collection and does not offer the user a pass/fail indication.

For the Zigzag tracking, the Logitech pen automatically records the user's time from start to finish, avoiding the need for manual timing with a stopwatch. The error score is based on a pixel-by-pixel comparison of the pen's position with the printed track and obstacles. Thus higher scores are given for greater distances penetrated into the obstacle. This allows a much greater precision than with manual scoring.

The trails test calculates the response times for each correct response and the order the user completed the test. In conventional pen and paper trails tests it is only usually possible to record the total time from when a subject started the test to when he/she finished. With the digital pen it is possible to record response times for each element of the trail.

RESULTS AND ANALYSIS

The applications were evaluated in a pilot study of 49 males and 18 females. Impaired and unimpaired data was collected from the University student bar, with the majority of normative data being collected at the Police Scientific Development Branch. Every subject was asked to perform each test once and then complete a breath test. The mean age for subjects was 26, with the standard deviation of 9.

Existing research highlights the impairing effect alcohol has on speed and accuracy. The maze test proved sensitive to these effects. The mean error score achieved by Birmingham subjects before consuming alcohol was 376 with a mean time of 40599 m/s. After consuming alcohol, the mean score was measured to be 633, with the mean time increasing to 47149 m/s. However, this does not imply that the test works perfectly well. Analysis of the frequency distribution indicates that the test gives 29% discrimination at the 77% level. This means that we cannot rely on the test to give us accurate results in discriminating between impaired and unimpaired users in an absolute sense, and hence its use in the field is limited. It can distinguish between impaired and unimpaired states for an individual if we have a baseline measure, however. These results are based on a small sample size that makes it unwise to use this data to make a general conclusion about the discrimination of the tests on the population; rather, it gives an indication of whether or not the test works in general.

From the trail it was evident that there was a significant difference in strategy by different age groups. It seems that the 18-25 sample falls into the quick and inaccurate category. As age increases the subjects become more cautions and accuracy improves whist time increases. This is quite interesting when you consider the current statistics for traffic accidents, with young drivers attributing to the highest percentage of RTA's. This may be due to inexperience, but could also be attributed to an increased degree of confidence regarding their skill level. The fact that our results indicate the 18-25 category sacrificed accuracy for speed helps to illustrate this point.

As well as age there was also a noticeable difference in scores been different genders. The trail demonstrated that males gained both significantly lower error and times than females. Computer games are predominantly played by young males and require skills similar to that displayed in the maze test. Games require the users to react quickly to changing environments, develop hand eye coordination and increase concentration spans. Ruff et al. (1993) confirmed the theory that there is a difference in hand eye coordination between genders. These factors could be why males scored significantly better both in both categories.

Using the methodology discussed by Gibson H. B. (1978), a scatter graph with the predicted regressions lines of time and error was created to demonstrate the different categories subjects fell into (see fig. 6). It can be observed that the majority of the impaired subjects (19 out of 27) fell within the quick and careless category, with a mean BAC level of 48 and a SD of 19. The remaining 8 subjects fell within the quick and accurate section, but had a lower mean BAC of 38 and a SD of 18. This finding is interesting in view of suggestions that alcohol may influence the speed-accuracy trade-off, leading to increased error rates (Tiplady, B et al. (2003)).

When tested repeatedly on a small sample of subjects there was no significant improvement in score, illustrating the fact that there is a minimal training effect. The ability to randomize the objects within the maze will further reduce any training effects.

The volunteers used in all trials were in a state of good health but lacked any people in the 65+ age group. Therefore, this will give an understatement of the variability of performance of the population. For a roadside testing device, the performance limit would probably be set at or around the 95% level thus incurring 5% false positives. The data presented in this project indicates that only a fraction of offenders will be identified with this test if the performance level is set at 95%. For a test to be more discriminatory, the mean scores of unimpaired and impaired performances must be as far apart as possible. This must be the main objective of future work on this test, and on the project in general. To make sure no training effect is occurring variations on the maze could be computed, with differing obstacle locations. There is also discussion of increasing the amount of data collected to see if there is any significance in how a subject incurs errors, either though the amount of times a subject leaves the track or how many times on object is glanced or penetrated.

Zigzag tracking has proven to be a very effective test in demonstrating the potential of the pen. The test requires no previous experience, no literacy or numeracy skills and is quick and easy to administer. With a relatively small sample it has been hard to find a correlation between score and time, but it is hoped that as data is collected a more appropriate scoring methodology can be established.

The trails test has not been as successful when used in conjunction with the digital pen. On occasions, successive, fast strikes made onto the page are not always registered. Although when tested with the Logitech IO software the pen appears to perform correctly, when used in conjunction with the SDK there are occasions when pen strokes are not detected. Whilst a sampling frequency of 100Hz may appear adequate, in practice this may not be sufficient for the fastest strokes that can be made

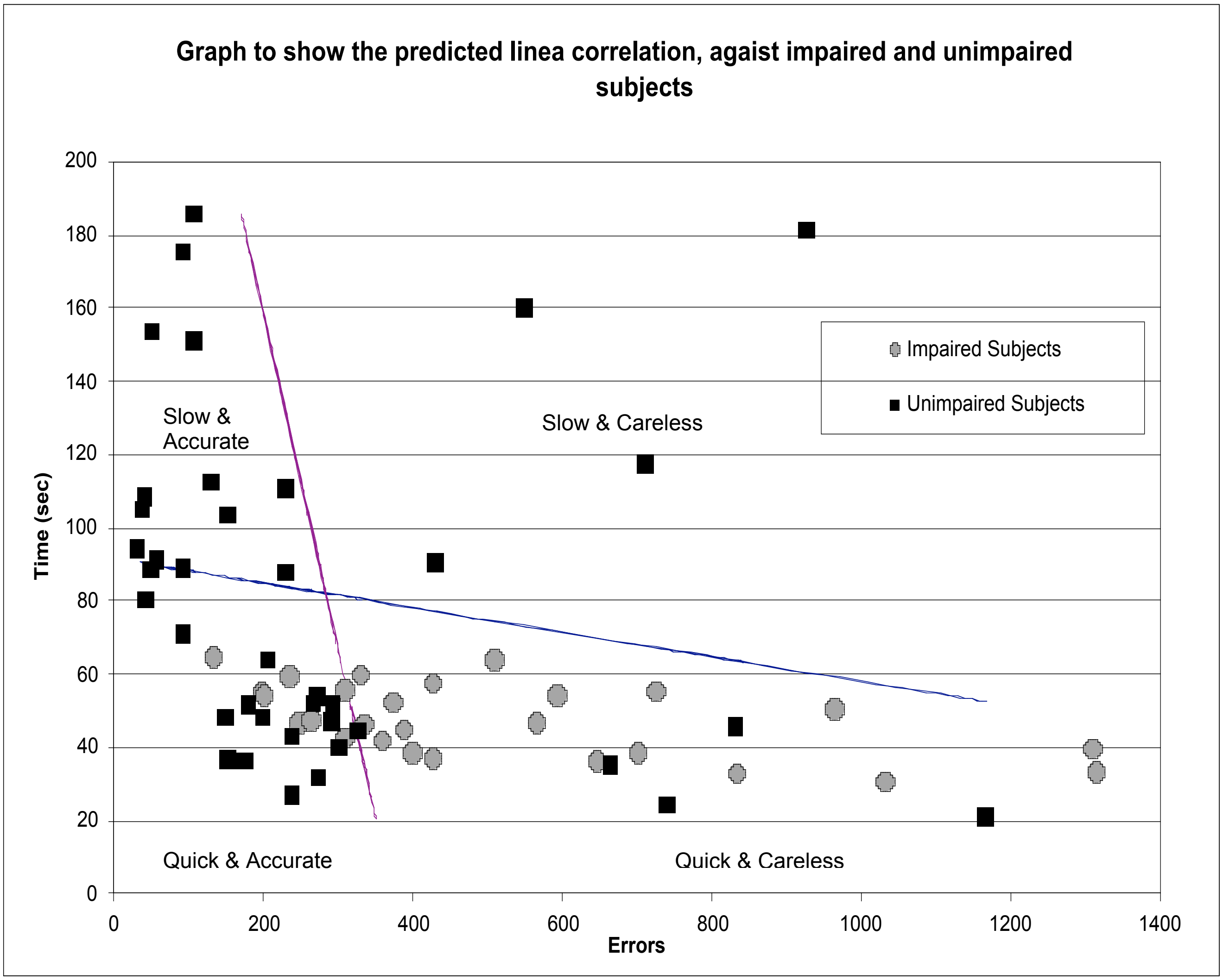

Figure 6: Scatter Graph of Impaired and Unimpaired Subjects

We also examined the pressure information received. The digital pen is capable of recording 125 discrete levels of pressure applied to the paper. However, the pressure exerted on the page by most subjects was at the top of this range so little additional information was gained via these pressure measurements.

One of the key elements of the trial was to validate the assumption that a pen and paper based test was highly acceptable to the users, and so a questionnaire survey of participants (n=67) in the preliminary trials was conducted. Seventy-three percent found the pen comfortable to use, with seventy-seven percent of experienced PDA users voicing the opinion that they would prefer to be tested with the digital pen, than with a stylus and PDA. The general consensus from the public and police officers interviewed was very positive and as the technology improves and reduces in size, general acceptance of the pen and associated tests should increase. From a usability perspective, the approach was very successful; users found the tests easy to understand and to administer, and the familiar nature of the devices meant that they had no issues in completing the tasks.

CONCLUSIONS

The Anoto pen offers a very attractive addition to the current technologies employed to test impairment. Improving the ability for users to complete a test battery without their score being biased by technological experience offers a much fairer, readily accessible test procedure. From a Police perspective the pen offers a light, reliable, relatively inexpensive device that may be developed with the additional Bluetooth functionally to offer a very effective roadside tool. Whether the digital pen could be used as an operational screening device by the police depends on the discrimination power of the actual tests and this aspect requires further investigation. Currently, we can determine impairment if we have a baseline unimpaired measure for an individual, but cannot with any accuracy identify impairment in the general population.

Possible Further Steps

This project has proven the potential for roadside testing with the digital pen. The next stage of the project hopes to increase the number of subjects studied, through both a controlled alcohol study and the evaluation against registered drug users. It is also necessary, as more data is gathered to improve the scoring system, combining time and errors in a more effective manner. Additional software add-ons are planned and it is hoped that a symbol substitution test can be computed to run alongside the current test battery, aiding the generation of meaningful absolute discrimination abilities that would allow this to be used as a real-world, effective test.

REFERENCES

Cameron, E., Sinclair, W., Tiplady, B. (2001) Validity and sensitivity of a pen computer battery of performance tests. Journal of Psychopharmacology 15: 105-110

Clark A & Riedel W. (2000). Roadside Impairment Testing Methods. CERTIFIED project report DR2a funded by EC under 4th framework programme.

Everest, Tunbridge & Widdop. (1998). The Incidence of Drugs in Road Accident Fatalities. TRRL Report RR202. Transport and Road Research Laboratory, Crowthorne.

Gibson H. B. (1978). Manual to the Gibson Spiral Maze 2nd edition Hodder & Stoughton Educational, Sevenoaks, Kent.

Hope, A., Woolman, P.S, Gray, W., Asbury, A.J & Millar, K. (1998). A computerised system for psychomotor evaluation: design, implementation and practice effects. Anesthesia 53, 545-550

O'Keefe, M. (2001). Drugs driving – standardised field sobriety tests: a survey of police surgeons in Strathclyde. Journal of Clinical Forensic Medicine. 8, 57-65

Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. Perceptual and Motor Skills, 8: 271-276.

Ruff R.M., and Parker S.B. (1993). Gender and age-specific changes in motor speed and eye-hand coordination in adults: normative values for the finger tapping and grooved pegboard tests.

Tiplady, B., Hiroz, J., Holmes, L., & Drummond, G. (2003). Errors in performance testing: a comparison of ethanol and temazepam. J.Psychopharmacol., 17: 41-49

Tunbridge & ROWE. (1998). The Incidence of Drugs in Road Accident Fatalities in Great Britain, 9th International Conference on Road Safety in Europe, Bergisch Gladbach, Cologne, September 1998.

White, A. M; Matthews, D. B. & Best, P. J. (2000). Ethanol, Memory, and Hippocampal Function: A Review of Recent Findings. Published by Wiley-Liss.