Human Factors Affecting the Control and Perception of Motor Vehicle Dynamics

A thesis

submitted as partial fulfilment

of the requirements for the Degree

of

Doctor of Philosophy in Psychology

at the University of Waikato

by

Kerry Spackman

University of Waikato

1996

Abstract

Vehicle Dynamics Engineers (VDEs) responsible for developing passenger vehicles and Elite Race Drivers were analysed over a 5 year period using telemetry data. Investigations were undertaken in an attempt to identify the most important rate limiting factors for each profession and to develop effective training methods for them.

In a preliminary investigation it was found that the VDEs were unable to reliably detect certain changes to a vehicle's dynamic behaviour despite their high confidence that they were doing so. This result led to the main VDE study where the VDEs were assessed using a Rotary vestibular platform, the RAF Basic Attributes tests and telemetry analysis of their performance while they evaluated changes to a vehicle's dynamic behaviour. It was found that there was no significant relation between either the Attributes scores or vestibular performance and the VDEs' ability to discriminate changes to the dynamic behaviour of a vehicle, termed their Evaluation performance.

After the VDE's baseline Evaluation performance was established, they received individual training to improve both their performance at sensing dynamic motion (Perceptual performance) and their consistency and accuracy of controlling vehicle motion (Control performance). The VDEs were then retested to determine the effect of this training on their Evaluation performance. Whereas before training most VDEs scored in the worst Evaluation performance category, after training most scored in the highest Evaluation performance category. A new mathematical technique for analysing telemetry data was developed which allowed us to separately determine the effects of Perceptual performance and Control performance on Evaluation performance. It was found that improving the VDE's Perceptual performance accounted for all improvements in their Evaluation performance whereas Control performance was unrelated to their Evaluation performance. This finding led to the development of a new Human Factors training program which will be instituted for all Ford's professional VDEs from 1996 onward.

In the Elite race driver study, a mathematical model was developed which allowed us to simulate the effect of altering a driver's curvature control strategy, which is a key component of a driver's style, on lap times. It was found that all drivers tested, with the notable exception of 3 times World Formula 1 Champion Jackie Stewart, had inadequate and / or faulty cognitive models for optimising curvature control when compared to our computer simulations. Training the subjects to improve their curvature control model produced substantial objective improvements in performance whereas additional unaided practice produced no improvement. For example a 19 year old N.Z. race driver was trained in one afternoon to lap faster than the current World Touring Car champion using the same vehicle.

Acknowledgment

I would like to acknowledge the helpful suggestions and excellent academic guidance I received from my supervisors Dr John Perrone and Dr Robert Isler in preparing this thesis. Thanks also to Dr Murray Jorgensen for his assistance in developing the statistical algorithms and to Ford Motor Co. who provided the facilities and funding that permitted this investigation. Finally to Dr Joanna Koppens for her study companionship during the many long hours of analysis and preparation.

Contents

Page

Page

Page

Page

Page

Attachment: Elite Race Driver Analysis and Training

List of Figures

Page

Page

Page

Page

List of Tables

Section A

Preliminary

VDE

Studies

1 Introduction - The Role and Historical Training of the VDE

Designing a completely new vehicle takes approximately 3-4 years and typically costs between US$ 1- 3 billion (Ford Motor Co., personal communication). Much of this time is spent using sophisticated computer modelling to analyse the performance and design of each component of the vehicle and then subsequently the behaviour of the whole vehicle as the sum of its components. Once the design engineers are happy with the design of the vehicle and its behaviour in computer simulations, a number of prototypes are constructed by hand. Each prototype costs approximately US$ 1 million to construct (Ford Motor Co. personal communication).

It is the job of the Vehicle Dynamics Engineer, commonly referred to as the VDE, to drive these prototypes and to evaluate their handling characteristics using a variety of test manoeuvres (Ford Motor Co. confidential document). Different manoeuvres are used to evaluate different aspects of the vehicle's behaviour. As a result of these evaluations, the VDE will suggest changes to the design of the vehicle which they hope will increase the acceptance of the vehicle by the target consumer. The changes suggested by the VDE are then made to the prototype and the VDE evaluates the effect of these changes on the vehicle's behaviour by repeating the tests. In this fashion the vehicle is continually Modified in a development cycle for a further period of 1 -2 years until it is felt that the vehicle is ready for the market. This VDE development process therefore occupies approximately 20 - 30% of the time taken to produce a new vehicle and has a critical impact on the overall cost of development, the speed with which the manufacturer can get a new vehicle to the marketplace and the final quality of the vehicle.

In particular the VDE is charged with making sure that the vehicle;

i) Handles in a safe manner for the target customer

ii) Subjectively has a good 'handling feel' for the target customer

iii) Has good outright objective performance

iv) Has a comfortable and compliant ride

Meeting these objectives is not easy as they are often in conflict. For example, reducing a vehicle's understeer will often increase the outright performance of the vehicle and increase its subjective handling feel at low speeds by making the vehicle more responsive, but it may also make the vehicle less safe to drive at high speeds for an unskilled operator. The VDE must balance these conflicting requirements to find an optimal compromise solution for the target customer. If the target customer is an 18-25 year old performance driving enthusiast (e.g., Ford Mustang), then the target dynamic behaviour of the vehicle will be quite different compared to that for a 60+ year old target driver (e.g., Lincoln Town Car).

The prototype optimisation process is further complicated because for most dynamic problems that the VDE wishes to correct, there are a number of alternate solutions. Each of these different solutions will introduce subtle but different interactions with other dynamic behaviours of the vehicle that the VDE is also trying to optimise. For example, a particular handling problem might be overcome by;

i) Altering the roll bars

ii) Changing dampers

iii) Using different bushings

iv) Altering the weight distribution of the vehicle

v) Using different tyres

Each solution may cure the current problem equally effectively, but will also produce different effects on other dynamic behaviours of the vehicle. The task of the VDE is to find an alteration which cures the current problem with minimal detrimental impact on the other dynamic factors.

Thus, the VDE must have an intimate appreciation of how each component of the vehicle affects its dynamic performance. This is achieved by detecting the <u>fine structure</u> of the dynamic behaviour of the vehicle which is associated with each component. Each component introduces both gross changes to the vehicle's behaviour which may be similar to that produced by many different components and also subtle nuances within that gross dynamic behaviour that act as unique signatures for each component. The VDE uses these fine structure signatures to determine which component is likely to yield the greatest benefit when changed. Thus although the VDE is designing a vehicle for the average customer who has only a modest perception of the vehicle's dynamics, the VDE must have a much greater sensitivity to vehicle dynamics in order to determine which components need to be modified and in what fashion to achieve the best compromise for the target customer. In other words, the VDE must not only be able to determine **that** something is wrong, but they must have additional skills to determine **what** is wrong.

Because changes cannot be made to a prototype vehicle immediately, the VDE must be able to store in their long term memory all the subtle nuances of the dynamic behaviour of the vehicle so that they can discriminate between changes to this fine structure after considerable time has elapsed since the last test. The VDE must further be able to compare the behaviour of this vehicle against the behaviour of many competitors' vehicles that they have also driven over the preceding years.

As each modification to the prototype vehicle is quite time consuming, the greater the accuracy of the VDE in determining what changes need to be made, the quicker the prototype can be turned into a production vehicle. Making wrong or unnecessary changes will delay the launch of the vehicle which can cost millions of dollars per day in lost production revenue and may provide competitive manufacturers with a marketing advantage. Allowing poorly developed vehicles onto the marketplace is even worse as a bad vehicle can harm the reputation of the entire company and reduce sales of other vehicles produced by the company. Thus the performance of a VDE has a most substantial impact on the profitability of a vehicle manufacturing company and as such any process which can improve the performance of the VDE will be of great benefit to the vehicle company.

Prior to this series of studies, VDEs at Ford acquired their expertise over a long period of apprenticeship underneath an experienced VDE. There was no objective or independent assessment of the performance of either the existing experienced VDEs nor the new trainees. There was not even an agreed set of objective measures by which a VDEs performance **could** be measured. The VDEs simply rated their own performance informally and infrequently (Ford Motor Co. personal communication).

2 Comparison of Driving Styles - Amateurs and Professional VDEs

2.1 Background to Study

During April 1993, Ford Motor Company's Advanced Vehicle Engineering Department conducted a small study to compare the behaviour of four different vehicles during a standard evaluation procedure called the Lane Change manoeuvre (Ford Motor Co. personal communication). Ford was interested in how these different vehicles responded to a range of different drivers and therefore both Professional VDEs and untrained Amateur drivers were used in the test. This rare availability of instrumented data from both Amateurs and Professional VDEs under controlled experimental conditions allowed me to independently test the following hypotheses in an unauthorised and subsequent analysis of the data. That the VDEs would;

i) As a group have a uniform or standardised style of executing the Lane Change

ii) Individually exhibit greater inter-trial consistency than the Amateurs

2.2 Method

2.2.1 Subjects

Nine drivers were selected from 2 groups;

i) Five Experienced Professional VDEs who performed Lane Changes as part of their job description during the evaluation of vehicles. These drivers are referred to as Experts (labelled as drivers 1, 2, 6, 7, 9).

ii) Four drivers who had never performed the Lane Change manoeuvre before. These drivers are referred to as Amateurs (labelled as drivers 3, 4, 5, 8).

2.2.2 Equipment

Vehicle A 1993 Ford XR Cougar instrumented as per Ford Internal Instrumentation Policy (Ford Motor Co. confidential document).

Track The Double Lane Change Manouevre was laid out as per Ford Internal Test Procedure (Ford Motor Co. confidential document). This layout is shown below in Figure 1.

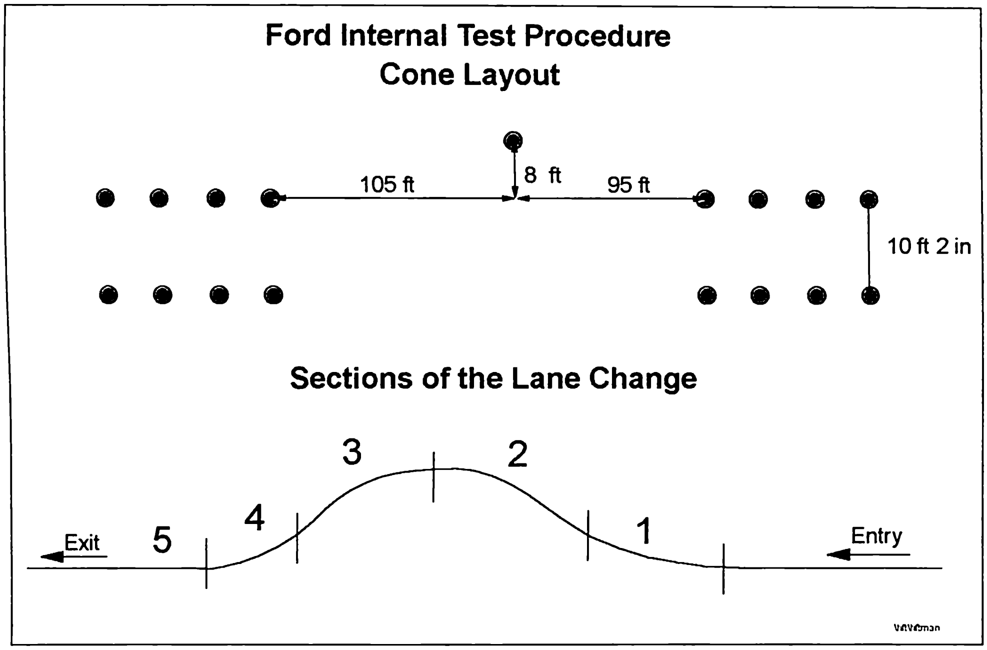


Figure 1 Layout of Lane Change

2.2.3 Procedure

For the Ford tests, each driver was required to drive 4 different vehicles through the Lane Change 15 times each at 45 m.p.h., after an initial practice session of 20 trials. For my analysis of the drivers, I utilised only the data generated during the trials with the 1993 Ford XR Cougar.

The test procedure for each driver was as follows;

a) 3 trials at 45 m.p.h. entry speed

b) 3 trials at 50 m.p.h. entry speed

c) 3 trials at 55 m.p.h. entry speed

d) 3 trials at 60 m.p.h. entry speed

e) 3 trials at 65 m.p.h. entry speed

The Experts were instructed to execute the Lane Change in the manner that they would normally use to evaluate a vehicle. The Amateurs were given no instruction other than to adhere to the correct entry speed and negotiate the cones.

2.3 Analytical Techniques

The inter-trial consistency and absolute driving style of each driver can be evaluated by defining metrics which characterise the driver's control inputs and the vehicle's dynamic behaviour output for each individual run. For example, the peak Steer Angle used by the driver during the entry section of each run is one such metric as is the mean Lateral Acceleration during the transition section of the Lane Change. The standard deviation of each metric for each driver is therefore a measure of that driver's control consistency while the mean value of each metric provides a measure of the driver's control style. Means and standard deviations were found at each speed for each driver and averaged within drivers to determine that driver's consistency and style at a given speed, rather than to measure how their style changed as the speeds changed. One hundred and sixty two different metrics were defined for each run, with each metric summarising a different aspect of the driver's control or the vehicle's behaviour.

One of the most important metrics for measuring how a vehicle responds to a driver's inputs in the Lane Change manoeuvre is the 'Yaw Hysteresis' metric (Ford Motor Co. confidential document). Yaw Hysteresis is calculated by plotting the Yaw Rate of the vehicle against the steering wheel angle (Steer Angle) and integrating the area contained within the curve. During normal driving, the Steer Angle, Yaw Rate and Lateral Acceleration of the vehicle are in phase. In this case the vehicle is said to behave linearly. However during a violent manoeuvre such as an accident avoidance manouevre where a driver swerves to avoid another vehicle, the Steer Angle is increased at such a rate that the vehicle cannot respond quickly enough and so the Yaw Rate is delayed relative to the Steer Angle with the Lateral Acceleration delayed even further. In this case the vehicle is said to behave non-linearly.

Figure 2 shows the Steer Angle, Lateral Acceleration and Yaw Rates during a normal driving manoeuvre while Figure 3 shows the same plots during a violent manoeuvre. In Figures 2 and 3 the Y-axis has been normalised so that under true steady state conditions the Yaw Rate divided by the Steer Angle and Lateral Acceleration divided by the Steer Angle, both yield unity. In figure 2 we see that the three curves are tightly grouped together while in figure 3 we see a marked separation of the curves in amplitude, time and shape.

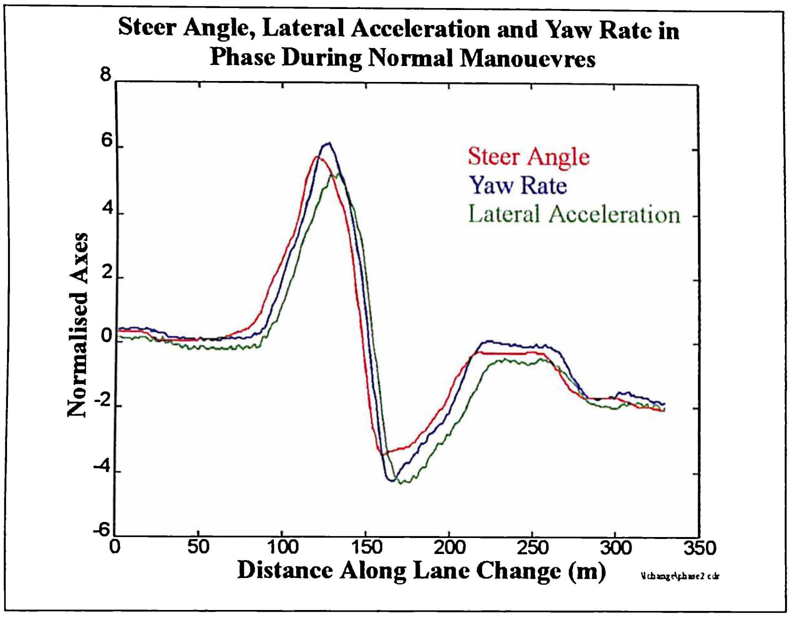


Figure 2

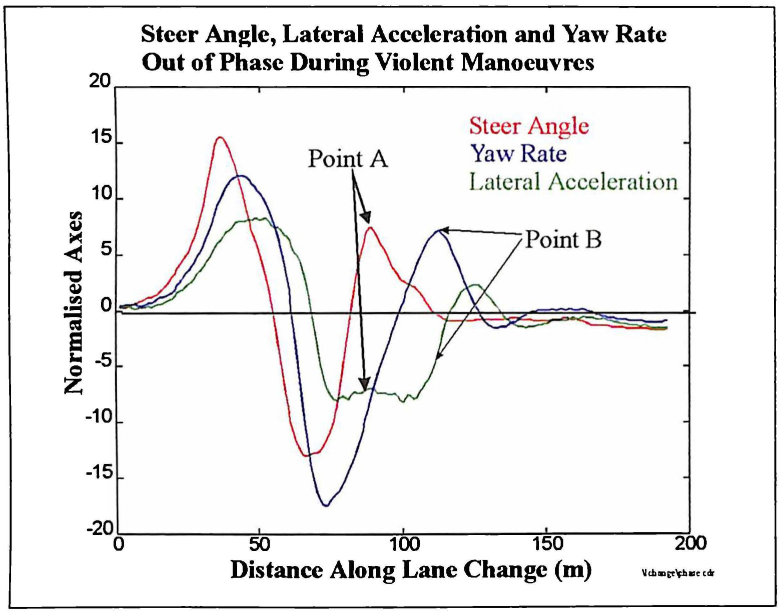


Figure 3

When the vehicle behaves linearly, the Steer Angle-Yaw Rate cross plot resembles a diagonal line because the two measures are in phase whereas when the vehicle behaves non-linearly, the Steer Angle-Yaw Rate cross plot contains large elliptical curves. By the above definitions, a vehicle which is behaving non-linearly will have a large Yaw

Hysteresis (because the curves contain a large internal area) whereas a vehicle which is behaving linearly will have a small Yaw Hysteresis. Figure 4 shows a comparison of the Yaw Rate -Steer Angle cross plots for linear and non-linear behaviour.

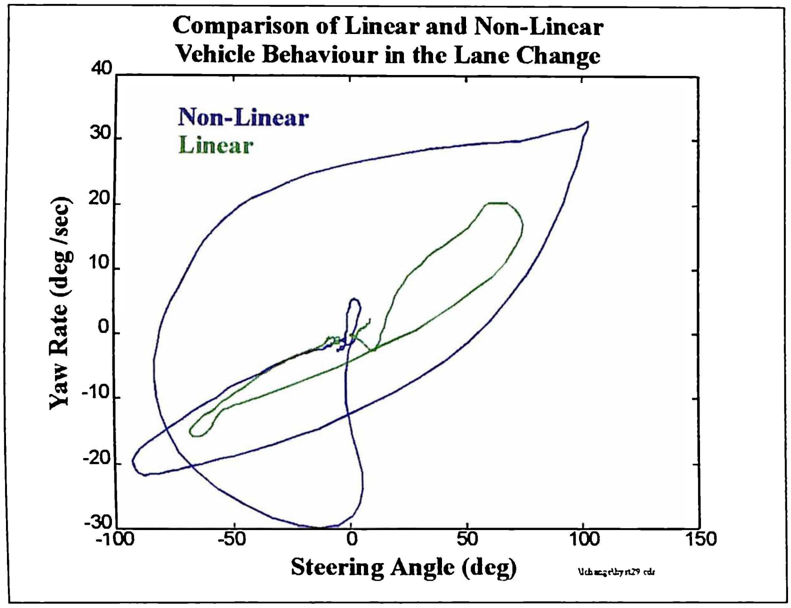


Figure 4

2.4 Results

2.4.1 Inter-Trial Driver Consistency

It was found that the Experts were no more consistent than the Amateurs when measured on each of the 162 different measures of inter-trial consistency. For example Figure 5 shows one of the measures, the mean Standard Deviation of the Peak Steer Angle used during the Entry section of the Lane Change, for each of the nine drivers.

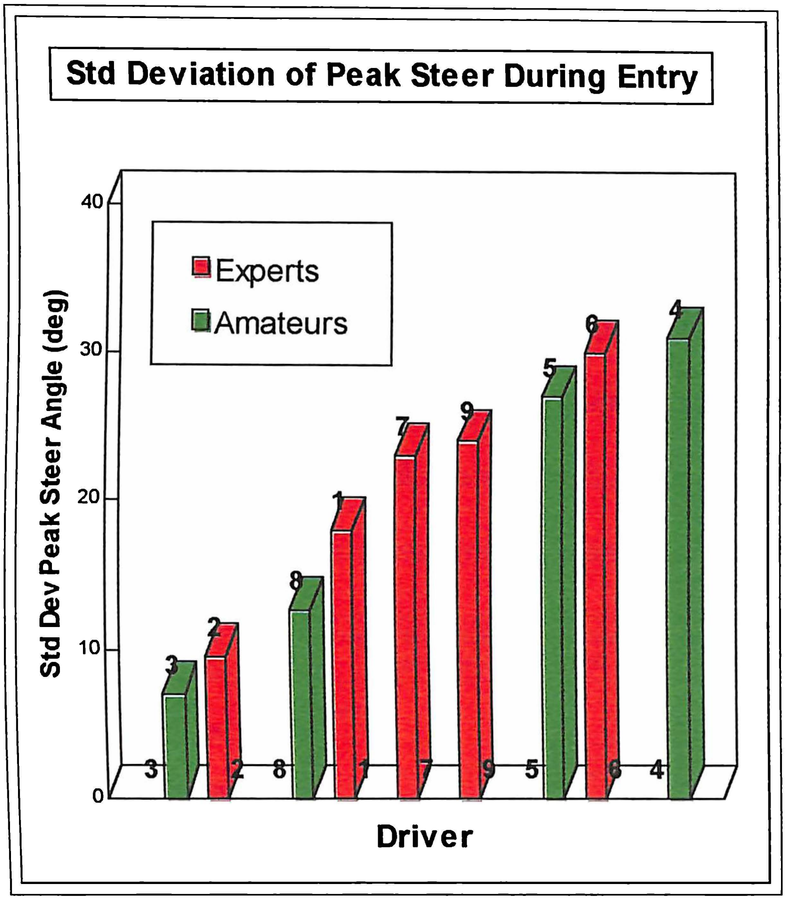
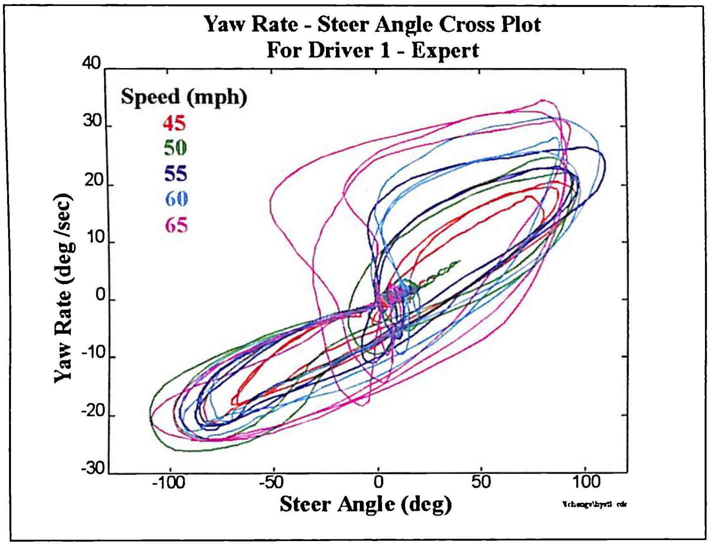


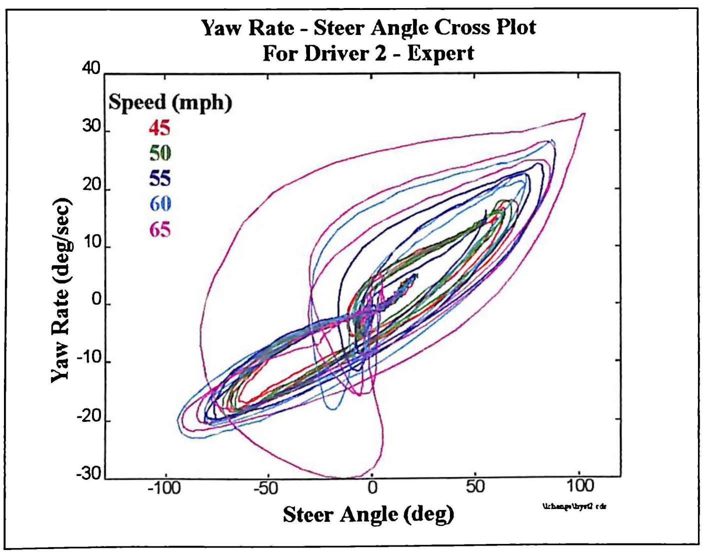
Figure 5

Figure 5 shows that it was not possible to separate the Experts from the Amateurs simply on the basis of their consistency. Furthermore, Figure 5 shows there was a wide variation in consistency between drivers with the most consistent drivers being approximately 4 times as consistent as the least consistent drivers. Both these aspects of Figure 5, namely the range of values exhibited by the different drivers and the inability to distinguish the Amateurs from the Experts by metric consistency, were typical of graphs obtained for all 162 control metrics.

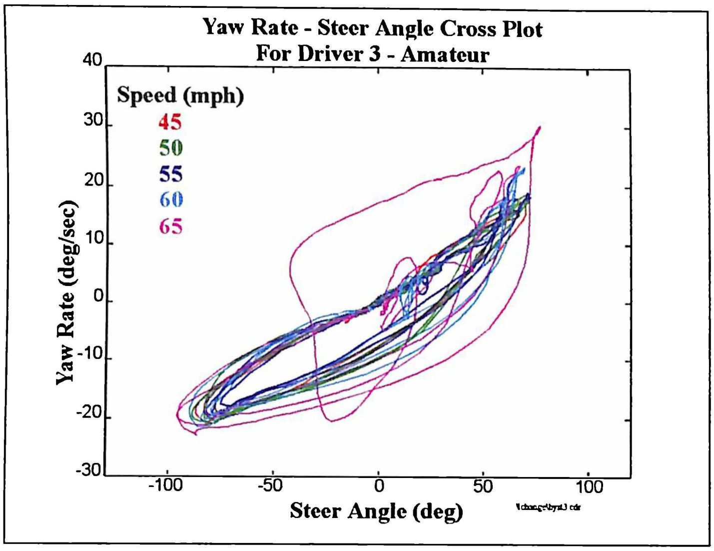
2.4.2 Absolute Driver Style

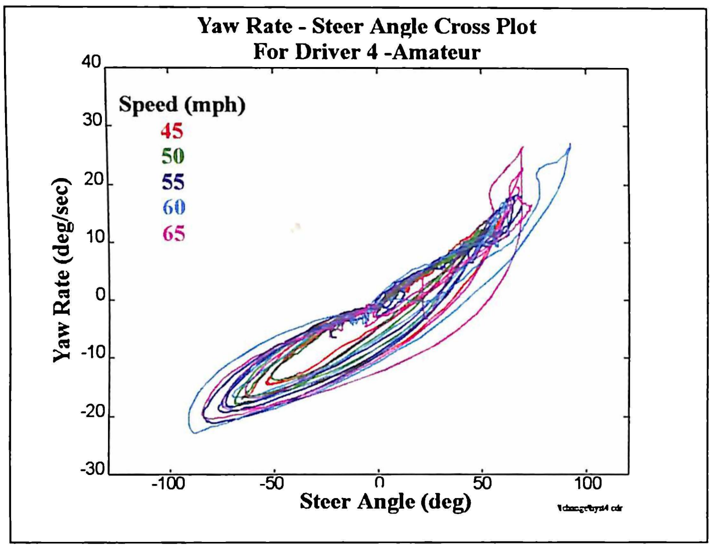
Figures 6 to 24 show that the Experts as a group did not have a uniform style of executing the Lane Change. Some Experts executed the manoeuvre in a very violent manner making the vehicle behave non-linearly, while other Experts executed the manoeuvre very smoothly making the vehicle behave in a linear fashion. It was not possible to identify the Professional VDEs from the Amateurs simply on the basis of their driving styles. Figures 6 to 14 below show the Steer Angle vs Yaw Rate cross plots for each driver for all their test runs (not including practice runs).

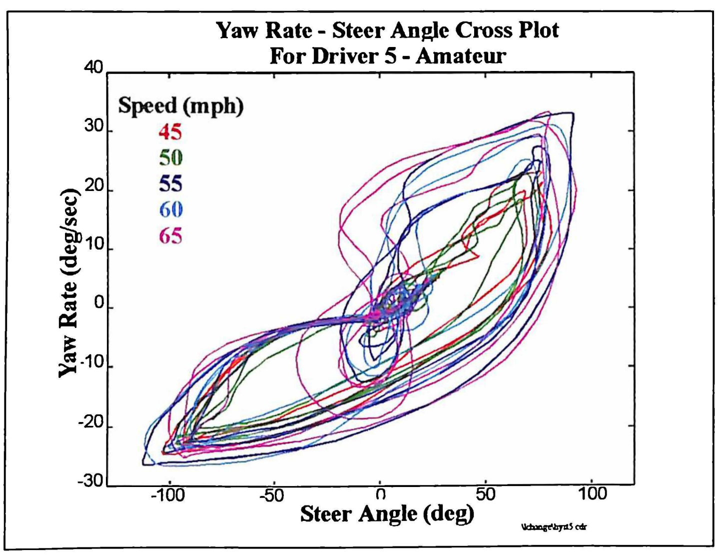




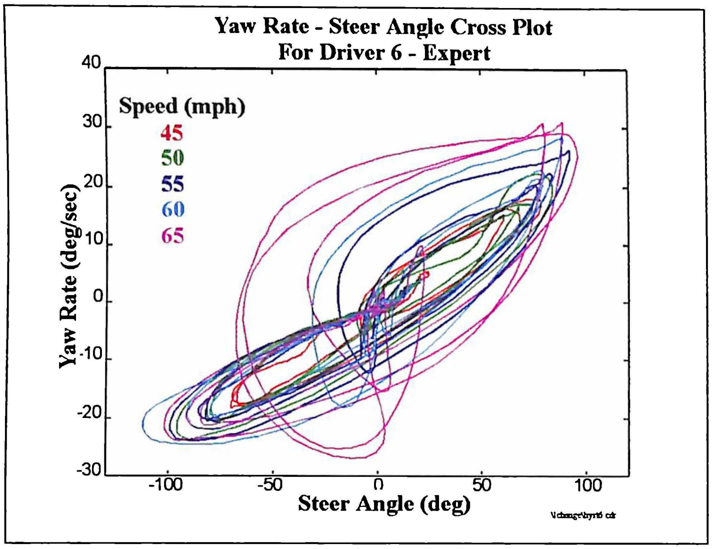
Figures 6 & 7

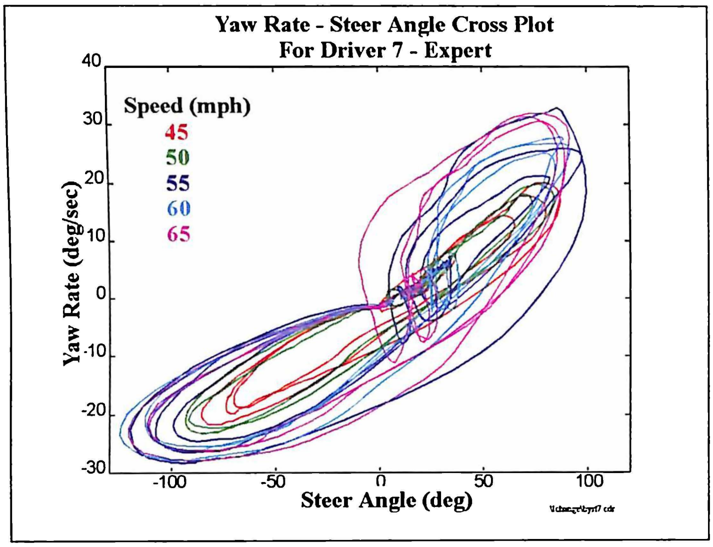

Yaw Rate - Steer Angle Cross Plot
For Driver 3 - Amateur
Speed (mph)
45
50
55
60
65
Yaw Rate (deg/sec)
40
30
20
10
0
-10
-20
-30
-100
-50
0
50
100
Steer Angle (deg)

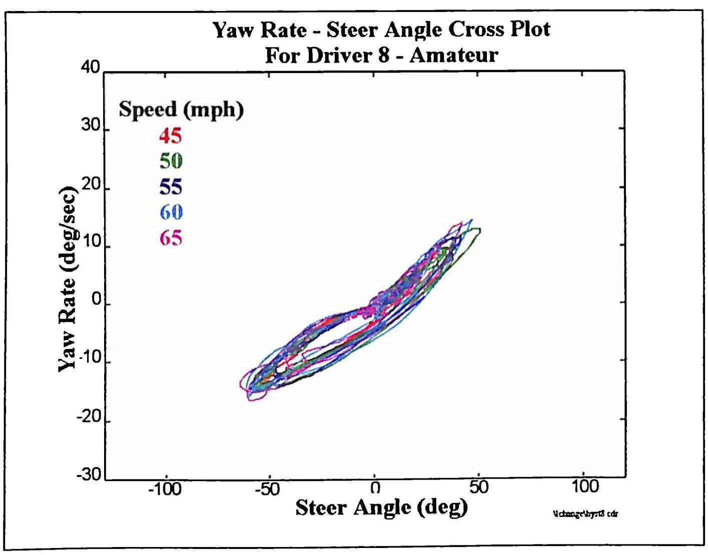

Yaw Rate - Steer Angle Cross Plot
For Driver 4 -Amateur
Speed (mph)
45
50
55
60
65
Yaw Rate (deg/sec)
40
30
20
10
0
-10
-20
-30
-100
-50
0
50
100
Steer Angle (deg)


Yaw Rate - Steer Angle Cross Plot
For Driver 5 - Amateur
Speed (mph)
45
50
55
60
65
Yaw Rate (deg/sec)
40
30
20
10
0
-10
-20
-30
-100
-50
0
50
100
Steer Angle (deg)

Figures 8 - 10







Figures 11-13

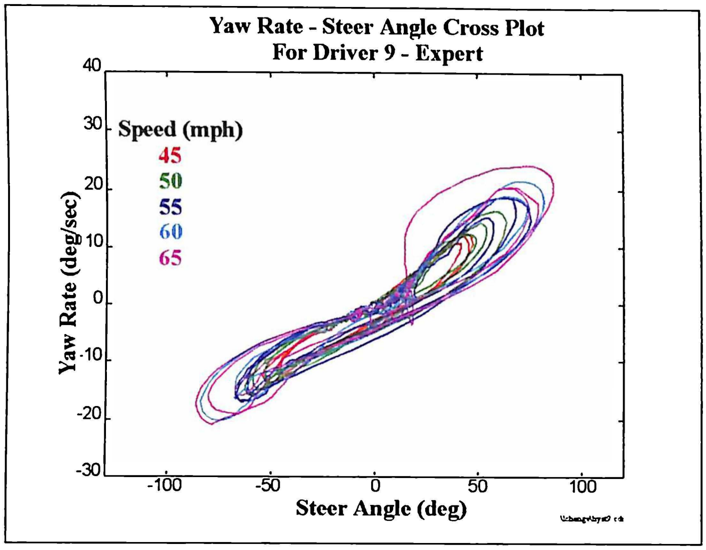


Figure 14

Figure 15 shows the total Yaw Hysteresis calculated from the above cross plots for all runs for each driver. The height of each bar in Figure 15 is equal to the total area contained within all the curves in the corresponding graphs of Figures 6-14.

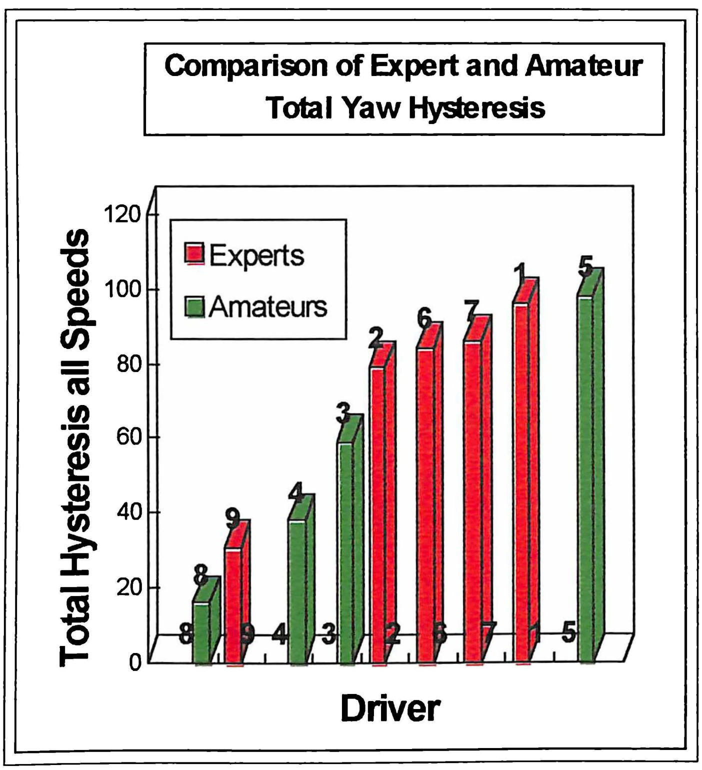


Figure 15

Figure 15 is a measure of the average non-linearity of the vehicle's response to each of the driver's inputs. Figures 16 to 24 below show the averaged steering inputs for each driver at each speed. The averaged traces were calculated by using fixed distance interpolation for all runs at each speed and filtering the mean interpolated values using a 2nd order Butterworth filter with cut-off frequency 0.02 (Oppenheim and Schafer, 1975; Rabiner and Gold, 1975).

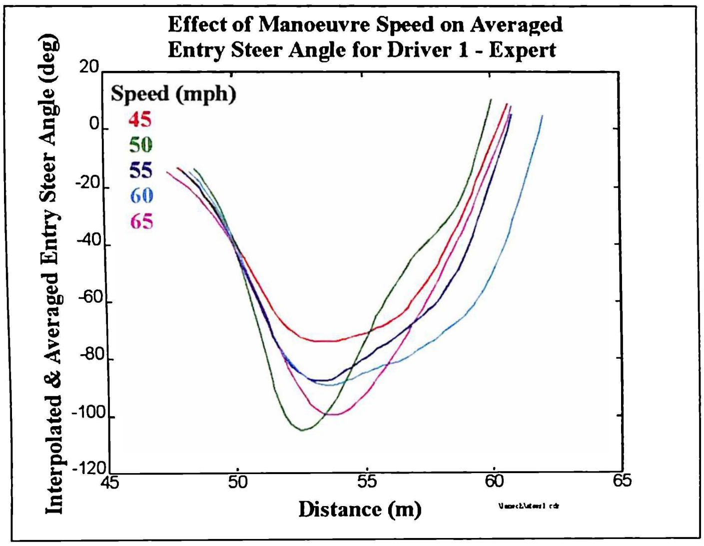


Figure 16

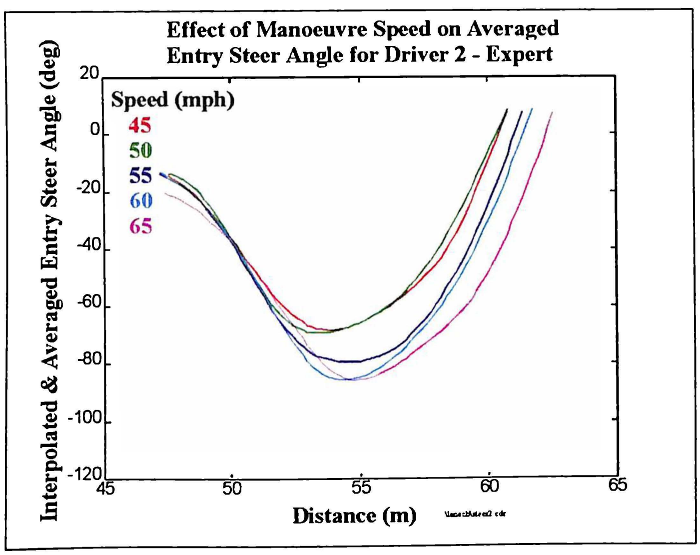
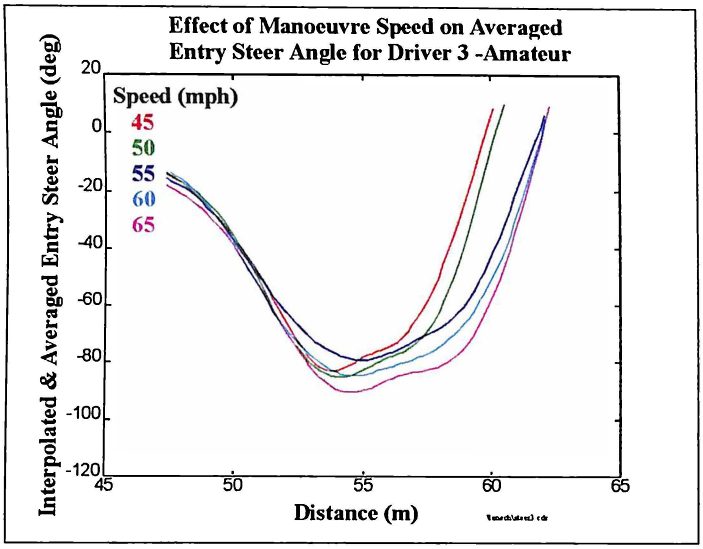
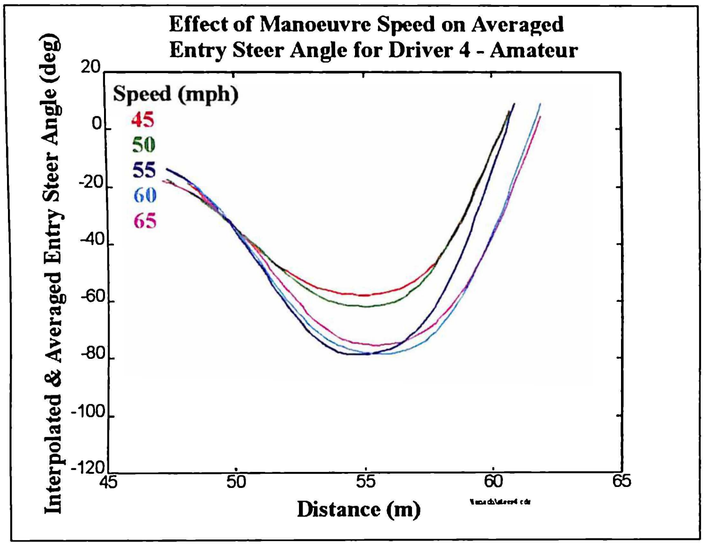
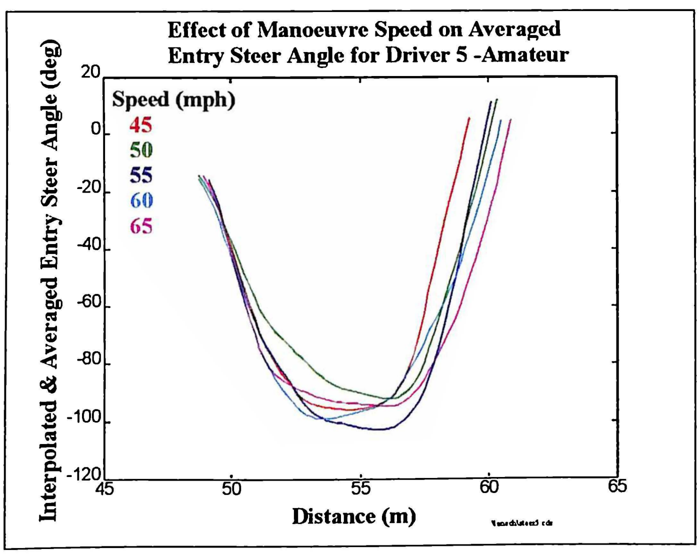


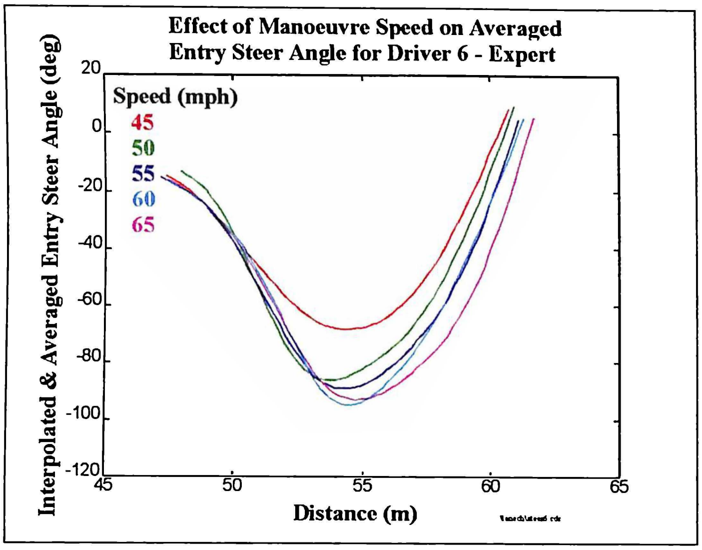
Figure 17

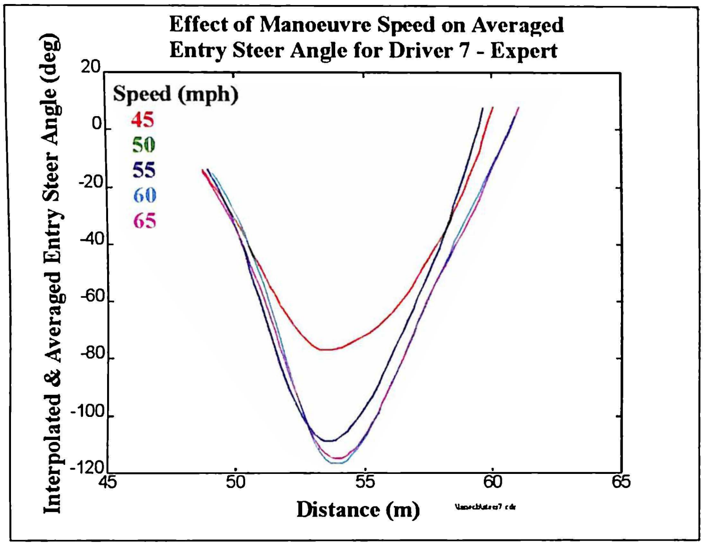


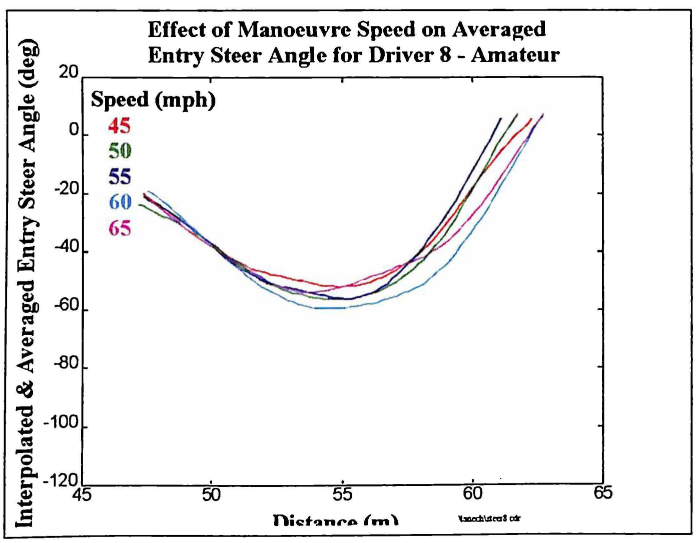




Figures 18 - 20







Figures 21-23

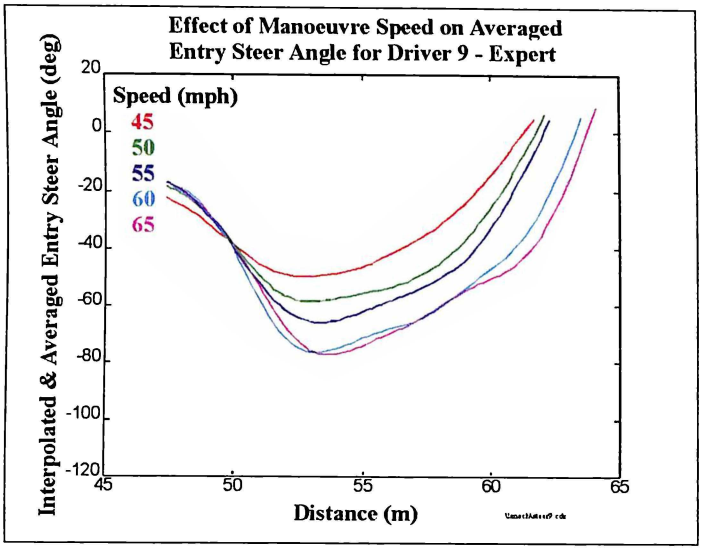


Figure 24

2.5 Discussion

VDEs perform the Lane Change manoeuvre to evaluate how safe a vehicle is during an accident avoidance event. In older design vehicles (pre 1980) it was often the case that when a driver swerved to avoid an obstacle, they would successfully miss the obstacle only to lose control of the vehicle and crash for example into an another oncoming car (Ford Motor Co. confidential document). The reason for the loss of control is explained by Figures 2 and 3 above. During normal driving the Steer Angle, Lateral Acceleration and the Yaw Rate are all closely in phase and the vehicle responds to the demands made by the driver (Figure 2). But in a highly non-linear manoeuvre the behaviour of the vehicle is significantly out of phase with the Steer Angle and so the driver must apply inputs which are not related to the sensations they are currently experiencing. Furthermore, some of these sensations will be in apparent conflict with others. For example, in Figure 3 at point B the vehicle is rotating in one direction but cornering in the opposite direction while at point A the driver is steering left and cornering to the right. Engineers try to design the vehicle so as to reduce this non-linearity and give normal drivers a better chance of controlling the vehicle in an accident avoidance situation. The Lane Change event is performed by the VDEs to assess how successful the design engineers have been in achieving this aim (Ford Motor Co. confidential document).

As we have seen, the non-linearity produced is a function of both the vehicle design and the method of execution. If the driver executes the Lane Change with very smooth inputs then the vehicle will display less non-linearity than if the driver applies violent inputs. The accepted conclusions at Ford Motor Co. arising from this dependence of dynamic linearity on VDE driving style were;

a) That the VDEs as a group must exhibit a uniform method of executing the Lane Change in order to achieve standardised rating of the safety of vehicles.

b) The VDEs must exhibit good inter-trial control consistency in order to reliably discriminate changes between suspension components in a single vehicle. (Ford Motor Co personal communication).

Figures 5 - 24 above show that the Expert VDEs as a group failed on criteria a) and most failed on criteria b). Because each VDE is assigned to develop their own vehicle and there is little if any interaction between VDEs during the development process across new models, then there exists the possibility that some VDEs may not be using the optimum driving style for evaluating vehicles. For example Driver 9 was the most experienced VDE at Ford and the one rated highest by his peers, yet Figure 15 shows that his driving style was approximately half as severe as the other VDEs. Driver 9 produced less than half the amount of Yaw Hysteresis as the next smoothest driver.

Further, many of the differences in vehicle behaviour brought about by suspension component changes are very small compared to the changes that result from different driving styles (Ford Motor Co. personal communication). If a driver is inconsistent then they may ascribe a perceived change in vehicle behaviour to a suspension change when it may in fact simply be the result of trial to trial changes in their driving style. If a suspension change is made to a prototype vehicle to reduce the amount of non-linearity by 10% but the driver's total non-linearity changes from trial to trial by 40% with the same suspension setting, then it is quite possible that the driver's inconsistencies will mask the vehicle modification changes.

I presented the above data to Senior Ford Management in August 1993 (Ford Motor Co. confidential document). This presentation coincided with serious difficulties in the development of the 1995 Ford Mustang. The VDE responsible for developing the Mustang had been making changes to the vehicle's Roll Bars over a period of 12 months and had accidentally arrived back at the original Roll Bar setting despite continually reporting improvements to the vehicle's handling with each successive new Roll Bar. My report and the difficulties with the Mustang program raised sufficient concern at Ford that the Vice President of Ford Motor Co. commissioned a preliminary investigation to further evaluate VDE performance.

The difficulty with the current analysis was that there was no independent measure of how sensitive the drivers were to changes to the vehicle's dynamics. The only measures available to us in this study were the VDEs' control consistency and control (driving) style. It could be that the VDEs were all still excellent evaluators of vehicle behaviour despite the differences in control consistency and control style exhibited by them. For example, the VDEs may somehow have been able to compensate for their inconsistency by remembering the differences in their inputs between successive trials and using some form of transformation to normalise their sensory experiences. Unless we could measure the VDEs' actual sensitivity to changes to a vehicle design we could not be sure how good they were as evaluators. It must be remembered that this study was not originally designed to evaluate VDEs but rather to study vehicle behaviour. The analysis of the VDEs was spontaneously undertaken after the event to satisfy my curiosity.

Section 3 below documents the exploratory study commissioned by the Vice President of Ford that arose from the current investigation.

3 Preliminary Investigation into VDE Discrimination Sensitivity and Confidence Ratings

3.1 Background to Study

After a VDE has performed a vehicle test, they will often make recommendations to senior management about modifications that they think are necessary to ensure a better product. These modifications can cost many millions of dollars (US$) and delay the introduction of a vehicle. Senior Management relies directly on the VDE's self reported confidence of their assessment as to the necessity of the modification. The greater the VDE's confidence in their own assessment, the more likely Ford is to spend money to modify the vehicle. It is therefore important to know both how sensitive a VDE is to changes made to a vehicle and how the VDE's confidence is related to their probability of being correct.

This investigation addressed these issues by determining;

a) Discrimination Sensitivities to suspension changes of Professional VDEs and Amateurs

b) How a Professional VDE's or an Amateur's probability of being correct in such a discrimination test is related to their self reported confidence that they are correct

The vehicle's Roll Bar stiffness was chosen as the adjustment that we wished to determine a VDE's sensitivity to, because this is a suspension adjustment that VDEs typically spend considerable time optimising. It was also the component that was incorrectly adjusted during the development of the 1995 Mustang.

The Roll Bar has two main functions. The primary function is to reduce the vehicle's body roll when the vehicle is cornering and the secondary function is to adjust the ratio of front tyre adhesion to rear tyre adhesion. Stiffening the front Roll Bar and softening the rear Roll Bar will make the front of the vehicle slide more than the rear of the vehicle during steady state hard cornering. Conversely, stiffening the rear Roll Bar and softening the front Roll Bar will make the rear of the vehicle slide more during a steady state corner. In addition to these normal gross changes to the vehicle's behaviour, adjusting the Roll Bars may produce other more complex secondary changes. When the vehicle undergoes a combined roll and pitch motion such as occurs in the Lane Change, the different suspension components are alternately compressed or extended. This alters the geometry of the suspension which in turn changes the angle at which each tyre is pressed onto the road and consequently how the cornering forces are transmitted through each tyre. This change in the vehicle's suspension position caused by roll and pitch therefore alters how much adhesion each individual tyre has. Each of the tyres will be affected differently, some gaining adhesion and some losing adhesion depending on the type of suspension used and the type of manoeuvre being performed. If the vehicle rolls more because the Roll Bar is softer, then each of the tyres will either gain or lose more cornering force than if the vehicle cornered flatter with a stiffer Roll Bar. This secondary change in adhesion caused by the change in the suspension geometry is added to the primary effect of altering front to rear adhesion ratios discussed above. During the Lane Change manoeuvre, the vehicle undergoes a very complex series of roll and pitch motions with many micro oscillations as the springs, dampers, Roll Bars, body torsional stiffness and bushes each contribute differently to the roll motion of the vehicle. In manoeuvres which are highly transient, such as the Lane Change where the vehicle is constantly changing direction and roll, the secondary effects of gaining or losing cornering force at each wheel caused by the suspension geometry change can sometimes overwhelm the primary effect of front to rear adhesion ratios. In these cases stiffening the front Roll Bar may actually **increase** the front tyre adhesion, contrary to expectation, as the roll is better controlled and the suspension geometry is not so adversely affected.

Changing the ratios of front to rear adhesion can have a considerable effect on the behaviour of the vehicle during the Lane Change manoeuvre because the amount of rotation or Yaw Rate of the vehicle depends critically on the lateral forces at the front and rear of the vehicle. Lots of front tyre slide will make the vehicle unresponsive during the initial entry phase where the driver is trying to initiate the turn and get the vehicle to begin rotating. Lots of rear tyre slide will make the vehicle oscillate on the exit and produce considerable instability called Yaw Rate overshoot. Thus, choosing the correct amount of stiffness for the front and rear Roll Bars is an important task as this has a strong effect on how the vehicle will behave and hence how the vehicle feels to the customer.

When assessing a vehicle the VDE must use both the primary and secondary cues to help determine whether a Roll Bar has been stiffened or softened.

3.2 Method

3.2.1 Subjects

Five subjects were selected by Ford Motor Co.

Subject D was an experienced male VDE specialising in Light Truck handling development. Subject D had worked extensively during the development of the Ford Explorer and had performed many thousands of Double Lane Changes.

Subject S was an experienced female brake test engineer specialising in Light Truck brake development. Subject S had a mechanical engineering background and was required to drive the vehicles to assess the development of new braking systems. While subject S did not normally perform the Lane Change manoeuvre during the course of her brake evaluation, she had previous informal experience with the manoeuvre.

Subject H was an experienced male VDE specialising in passenger vehicle handling development. Subject H had performed many thousands of Lane Change manoeuvres in vehicles. Subject H had been a project leader in the development of many vehicles.

Subject K was a male sound engineer responsible for instrumenting, testing and analysing the noise inside light trucks. Subject K had never tested a vehicle for any aspect of handling performance before.

Subject B was a male computer engineer responsible for designing and computer modelling new vehicles. Subject B had no previous experience with evaluating any aspect of handling performance.

There were thus 2 Professional VDEs who had specialised in the Lane Change Manoeuvre, 1 Professional VDE who had specialised in braking manoeuvres and 2 Computer Engineers with no vehicle evaluation experience. Only 5 subjects were available because of the cost of performing these tests (approx. US$ 230,000.00 for the five subjects) and the unavailability of the VDEs and the test track facilities.

3.2.2 Equipment

a) Vehicle and Instrumentation

A 1994 Ford Explorer (this was a pre-production model not yet available to the general public as the tests were conducted in early 1993) was used as a test vehicle and was modified by fitting specially fabricated adjustable front and rear Roll Bars.

The adjustable Roll Bar system provided 7 adjustment positions on either side of the front Roll Bar (total 14 adjustments) and 13 adjustment positions on either side of the rear Roll Bar (total of 26 adjustments). These adjustments could be made within 10 seconds without the driver's knowledge by two mechanics who waited on either side of the Lane Change start position. In addition to manufacturing adjustable Roll Bars for this experiment, demountable drop bar attachments were constructed so that the vehicle could be run with either or both of the Roll Bars disconnected. If the rear Roll Bar was disconnected then this would amplify any changes made to the vehicle's behaviour when the front Roll Bar was adjusted. For example, altering the front Roll Bar from setting 2 to setting 3 with the rear Roll Bar disconnected had a greater effect on the vehicle's behaviour than if the front Roll Bar was adjusted from setting 2 to setting 3 with the rear Roll Bar still attached. In this manner it was possible to amplify the changes to the vehicle's behaviour if the adjustments to the Roll Bars were too fine for the driver being tested to perceive.

During these tests, the left and right side of the Roll Bars were always altered together so that for example if the left side of the front Roll Bar was set to setting 3, then so would the right side of the front Roll Bar. In this case we would simply say that the front Roll Bar was in setting 3. During the Discrimination test session, it was possible to alter either or both Roll Bars depending on how hard we wanted to make the test. The front Roll Bar produced a greater change in vehicle behaviour than the rear Roll Bar, partly because it was thicker and shorter having greater torsional rigidity.

Thus we had 5 different Roll Bar adjustment possibilities for conducting the Discrimination test. They were in increasing order of difficulty;

i) Adjust front Roll Bar - rear Roll Bar disconnected

ii) Adjust rear Roll Bar - front Roll Bar disconnected

iii) Adjust front Roll Bar - rear Roll Bar connected

iv) Adjust rear Roll Bar - front Roll Bar connected

v) Adjust both front and rear Roll Bars, subject has to identify which of the 2 Roll Bars was adjusted for the Modified run **and** which run A or B was with both Bars in the Base condition.

The Explorer was fitted with a Cranfield RaceCorder data logger (Cranfield Institute of Technology) which had 12 bit resolution, 2 MB of RAM and 40 allocatable channels, 7 of which were digital. The logger was turned on and off automatically by infrared triggers which were positioned at the beginning and end of the Lane Change course. Data was recorded at 20 Hz sampling frequency.

The following telemetry sensors were fitted to the Explorer;

Lateral Accelerometer - Schaevitz +/- 5 g Servo

Longitudinal Accelerometer - Schaevitz +/- 5 g Servo

Yaw Rate Gyro - Systron Donner Solid State

Throttle Position Sensor (L.V.D.T.)

Steering Position Sensor (L.V.D.T.)

Brake Position Sensor (L.V.D.T.)

Brake Pedal Pressure Sensor (Strain Gauged)

Brake Line Pressure Sensor (Fluid psi)

Four Suspension Travel Position Sensor (L.V.D.T.)

Wheel Velocity Sensor (100 pulse per revolution encoder)

b) Test Track

Testing was conducted on the main straightaway at Ford Motor Company's Vehicle Evaluation Centre at Naples, Florida. The Lane Change layout used in this test was different to the standard Lane Change layout normally used at Ford (Figure 1). The reason for modifying the standard Lane Change layout was that during our experimental preparation in England, we found that the Roll Bar adjustments were difficult to detect using the normal Lane Change layout. We therefore spent 2 months preparation with Jackie Stewart and 5 Professional race drivers trying to make the adjustments produce a greater effect. This was achieved partly by the addition of the demountable drop bars discussed above and partly by the layout of the Lane Change. The modified Lane Change configuration is shown below in Figure 25.

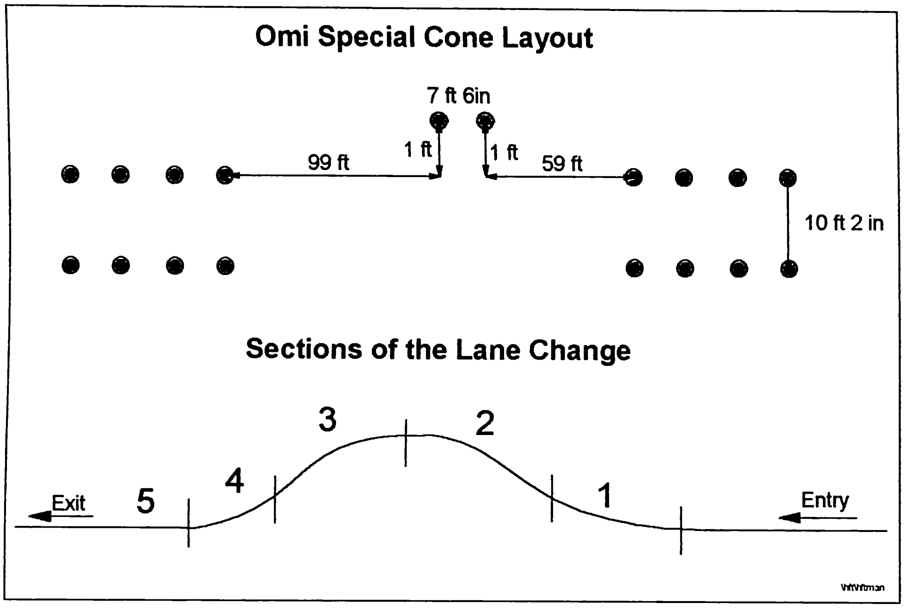


Figure 25 Modified Lane Change Layout

3.2.3 Procedure

a) Practice Session

The drivers were given 20 minutes of uninterrupted practice driving the Ford Explorer through the Lane Change with the Roll Bars being set in the Base condition (i.e., the softest setting number 1). During this Practice session, the drivers were instructed;

i) To become familiar with how the Explorer behaved in the Base condition so that in the later Discrimination test session they could discriminate between the Base and Modified settings

ii) To develop a driving style that was as consistent as possible so that variations in their driving style would not mask the effects of Roll Bar changes

iii) To drive the vehicle in such a fashion that it would tend to maximise the effect of changes to the Roll Bars

iv) To keep the Explorer as close to 75 km/h as possible throughout the Lane Change

b) Discrimination Test Session

Following the Practice session the drivers were tested in the Discrimination test. The Discrimination test consisted of repeated pairs of Lane Change manoeuvres (Run A and Run B). One of the pairs (for example Run B) would be with the Roll Bar set on the softest condition (Base condition) while the other run (in this case Run A) would be with the Roll Bar set on a stiffer Modified position. The drivers had to determine using a 2 alternate forced choice procedure which run, either A or B, was the Base run. The Roll Bar was adjusted using the method of constant stimuli. (Guilford, J.P. 1954). In addition, the subject had to report how confident they were in their choice using the 0 - 10 scale below;

0 - the driver has absolutely no idea at all, is simply guessing

3 - suspects the answer is right but not very certain

5 - reasonably sure the answer is right

7 - quite confident the answer is right

10- the driver is absolutely sure, there is no possibility at all that they could be wrong

All subjects were initially tested using the easiest Roll Bar test procedure (the front Roll Bar was adjusted with the rear Roll Bar detached) with the exception of Subject D who was initially tested on the hardest test procedure (adjustments made to both Roll Bars). This hard test procedure was found to be too difficult and therefore for all subsequent tests all subjects were tested on the easiest test procedure.

c) Informal Training

Each driver was given a series of informal Practice and Training sessions after their first Discrimination test session. The purpose of these sessions was to try to gain an initial appreciation of the factors that affected VDE discrimination performance. These investigations were necessarily informal because Ford Motor Co. viewed this current study as a pilot for a more formal study and therefore our brief was to quickly survey a range of factors that might be of interest. Thus different training procedures were used for different VDEs. Further, because both track time and VDE time was extremely difficult to obtain it was not possible in this pilot study to use control subjects that received no training. This topic is addressed in greater detail in Section 6.6 below.

d) Informal Training Procedures for each Subject

Subject H, the most experienced VDE, was administered the Discrimination test 3 times over a 2 day period. After each test he was told his results and then allowed to practice for as many runs as he liked. He was able to ask for any Roll Bar setting that he liked during these Practice sessions so that he could attempt to improve his discrimination performance. After these 3 sets of Discrimination test and Practice sessions, subject H was given telemetry feedback training to improve his throttle control. Telemetry feedback training was conducted in blocks of ten runs at a time. After each set of 10 runs, subject H was shown the telemetry results for those runs and weaknesses in his technique were identified. Targets were established for the telemetry traces and the subject was given a further set of 10 runs to attempt to achieve those targets. This feedback and training cycle was repeated a further 3 times.

Subject B was trained to be more consistent in his steering and throttle inputs by the use telemetry feedback training.

Subject D was trained to be more consistent in his steering inputs by the use of telemetry feedback training.

Subject K's discrimination sensitivity was re-measured at 70 km/h instead of 75 km/h.

Subject S was instructed to alter her method of analysing the vehicle motion by being told to concentrate on the roll motion of the vehicle instead of the understeer ratio.

3.3 Analytical Techniques

A subject's confidence ratings and accuracy can be plotted simultaneously for each stimulus level presented to them, by constructing a Confidence - Sensitivity graph. The x-axis plots the Modified Roll Bar setting for each trial pair and the y-axis plots the subject's confidence for that trial. If the subject is wrong in their choice the confidence value is multiplied by -1 to yield a negative integer. Figure 26 below shows a typical Confidence - Sensitivity graph produced by a good VDE.

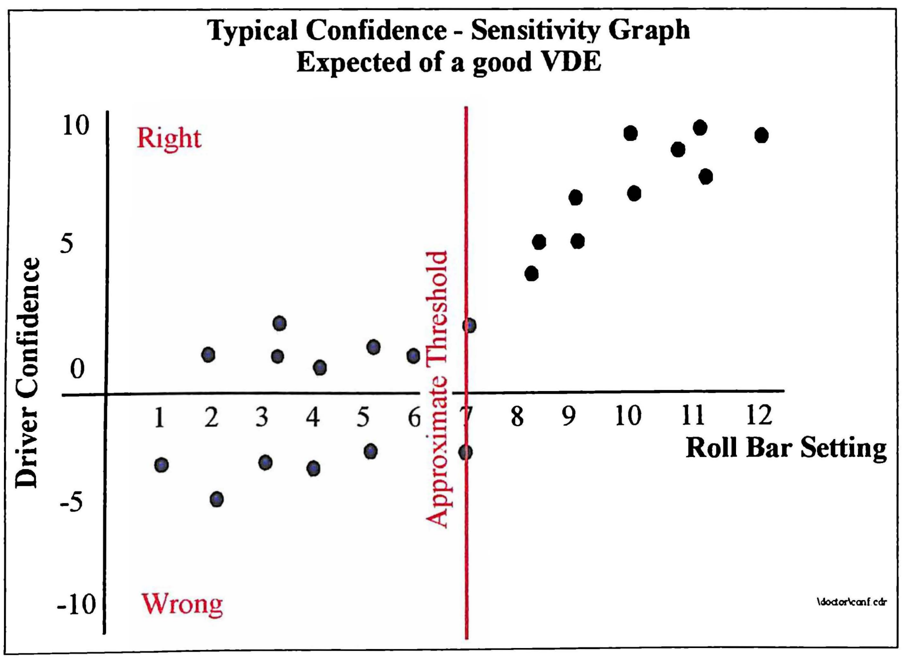


Figure 26

For stimuli pairs that are easy to discriminate (right side of graph), a good VDE would report correctly and with high confidence. As the pairs become more difficult (moving towards the left on the graph), a good VDE would respond with a lower confidence rating. Below the VDE's discrimination threshold, the VDE would respond with low confidence and a random probability of being correct.

3.4 Results

3.4.1 Initial Discrimination Test

i) All subjects scored at chance level on their first Discrimination test for all levels of stimuli presented.

ii) All subjects' confidence ratings were unrelated to their accuracy.

Confidence-Sensitivity graphs for the subjects' first Discrimination sessions are shown below in Figures 27-31.

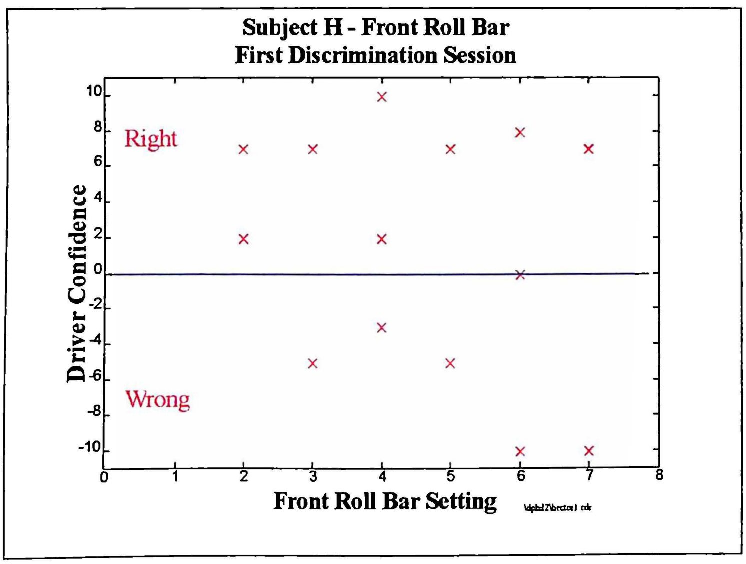
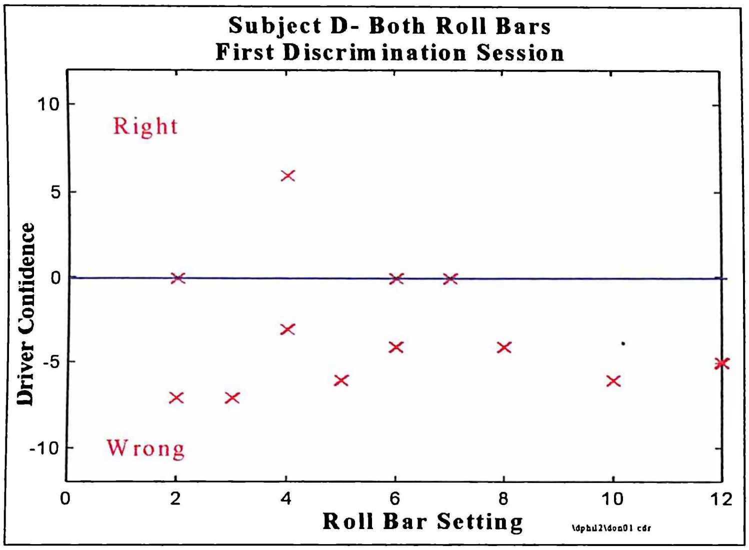
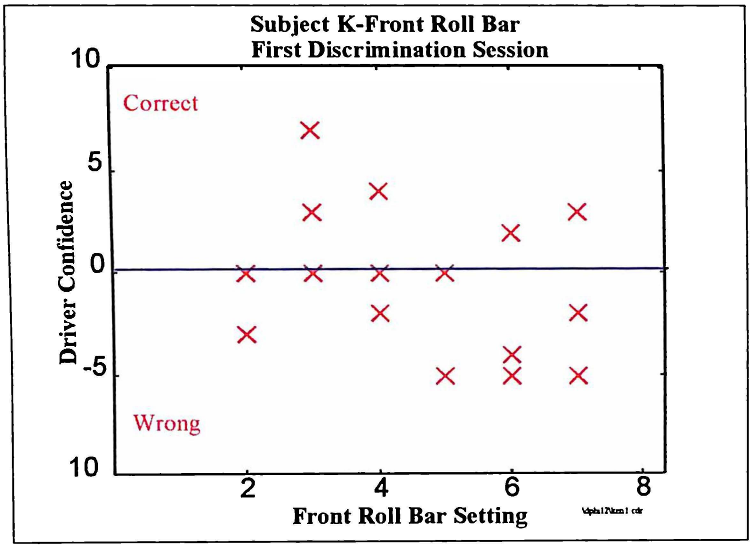
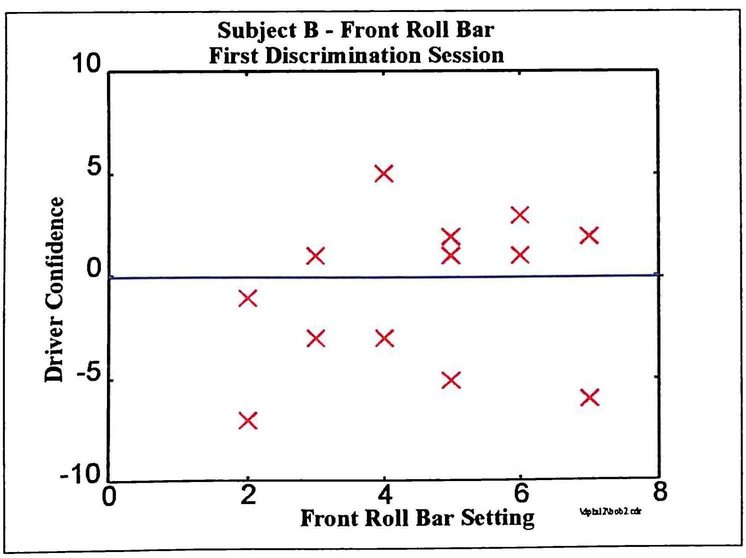


Figure 27







Figures 28- 30

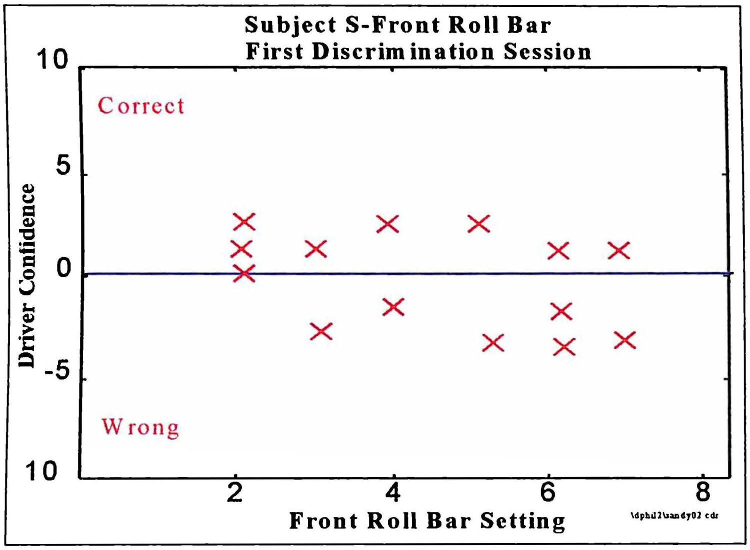


Figure 31

3.4.2 Informal Training

After more than 300 Lane Change practice trials over a 2 day period and 3 separate Discrimination sessions where he was told his results, the most experienced VDE subject H, still scored at chance level on all settings. Despite these poor results and his knowledge of failure, subject H had an average confidence rating of 7.2 and a most common (mode) rating of 10. Figure 32 below shows how subject H's confidence was related to his probability of being correct for all 3 Discrimination sessions combined.

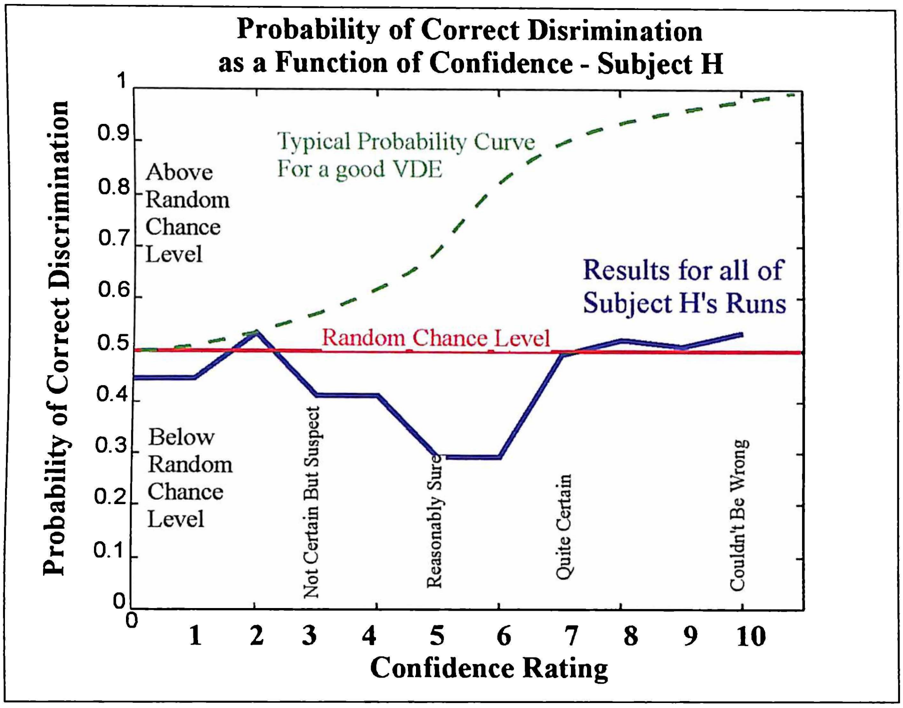
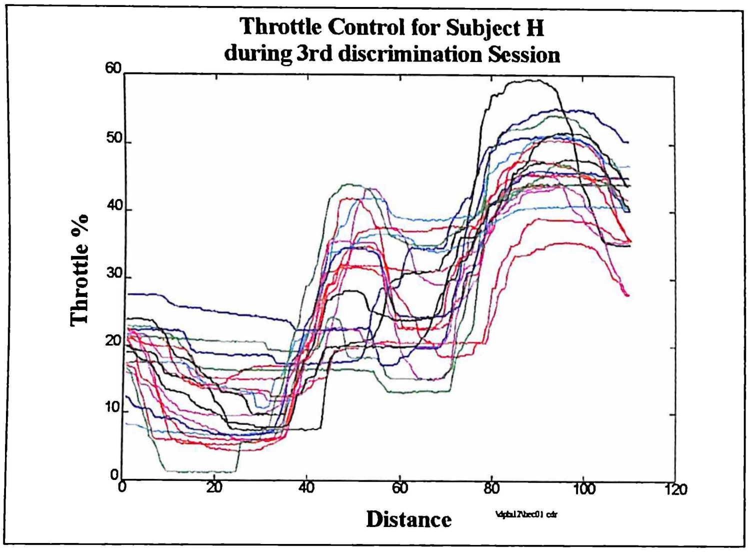
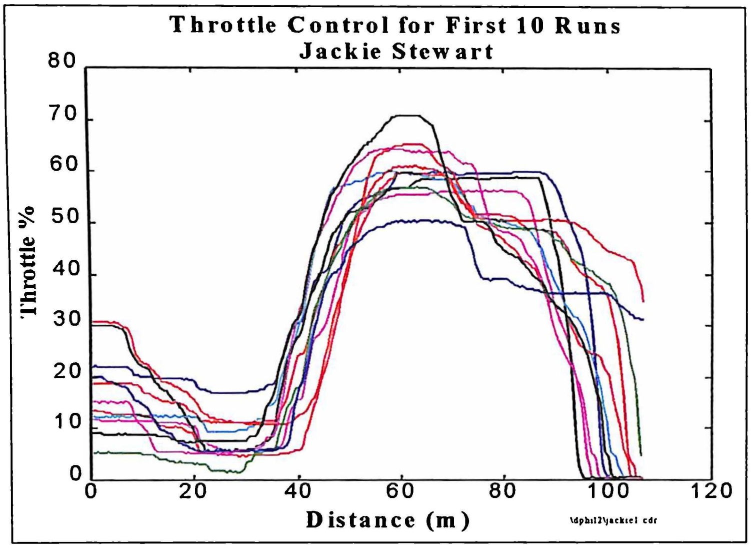


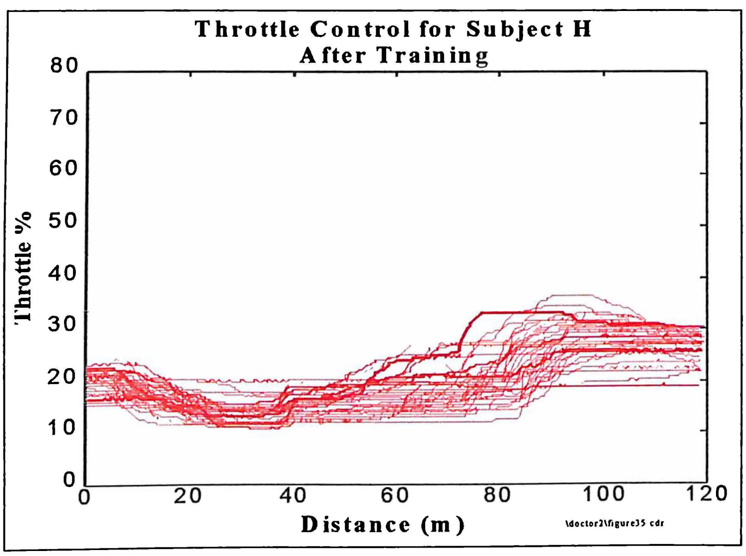
Figure 32

Figure 33 below shows subject H's throttle control during his third Discrimination session after approximately 300 practice runs had previously been completed. Figure 33 shows considerable scatter in his throttle applications between runs. Figure 33 can be compared to the throttle graphs from Jackie Stewart which were obtained during our preliminary testing in England prior to conducting this study. Jackie assisted us in our development of the Roll Bars and course layout as previously mentioned. Figure 34 shows Jackie Stewart's **first 10** practice runs in the Explorer. Figure 34 shows that Jackie immediately developed a highly consistent throttle style despite the large amplitudes of his throttle inputs.

Figure 35 shows subject H's throttle control after the throttle telemetry feedback training which shows a big improvement in consistency achieved primarily by reducing his inputs to a minimum.







Figures 33 - 35

Figure 36 below shows the results of feedback training on subject B's throttle and steering input consistencies. In this figure, driver consistency is calculated by finding the summed standard deviations at fixed distance interpolations for successive blocks of 10 steering and throttle curves. The consistencies were normalised so that the initial level on the first 10 practice runs was set to 100.

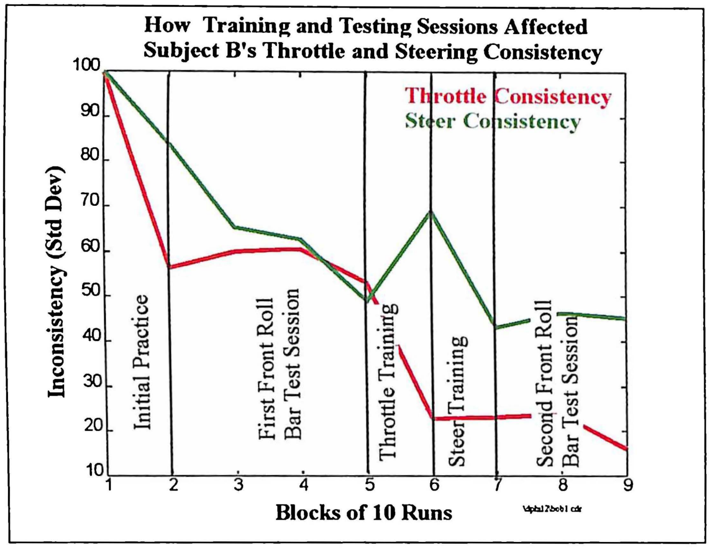


Figure 36

Figure 37 shows subject B's 2nd Discrimination test result after this training.

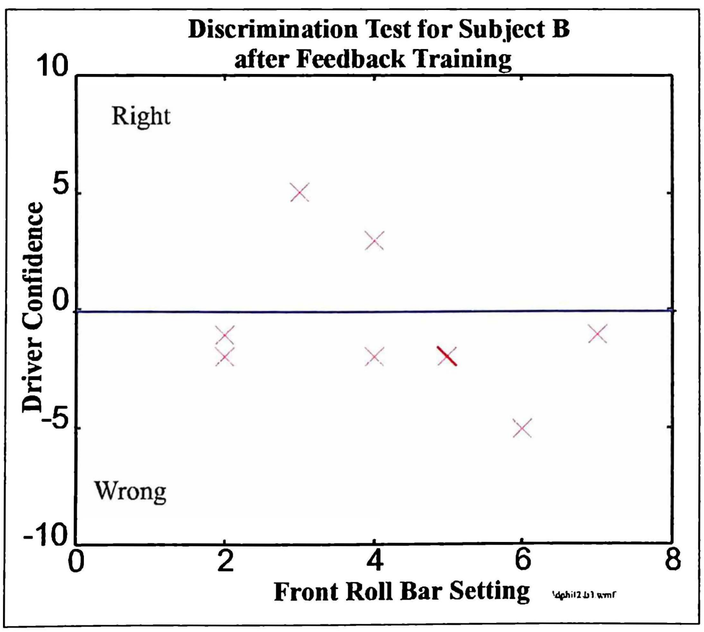


Figure 37

Figure 38 below shows the results of telemetry feedback training on Professional VDE subject D's steering consistency together with the steering consistency of subject S. Figure 39 shows the effect of this training on subject D's 2nd Discrimination test.

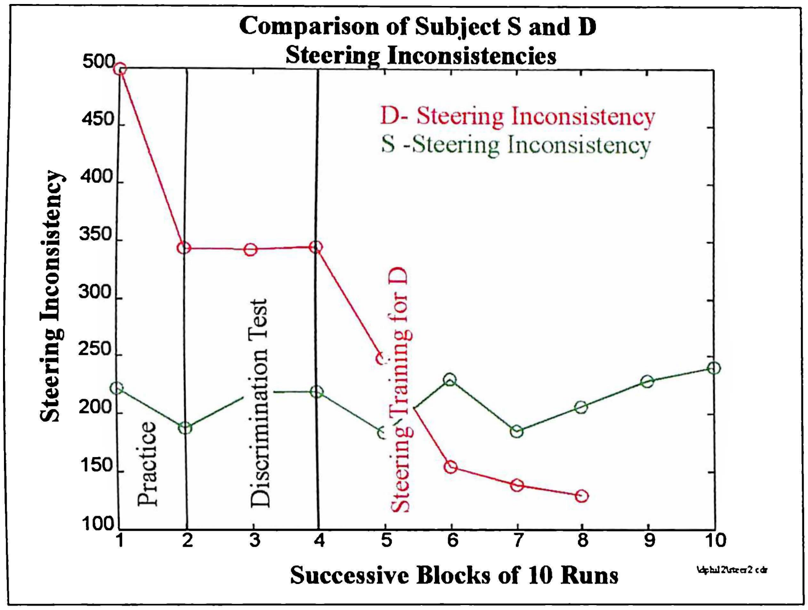


Figure 38

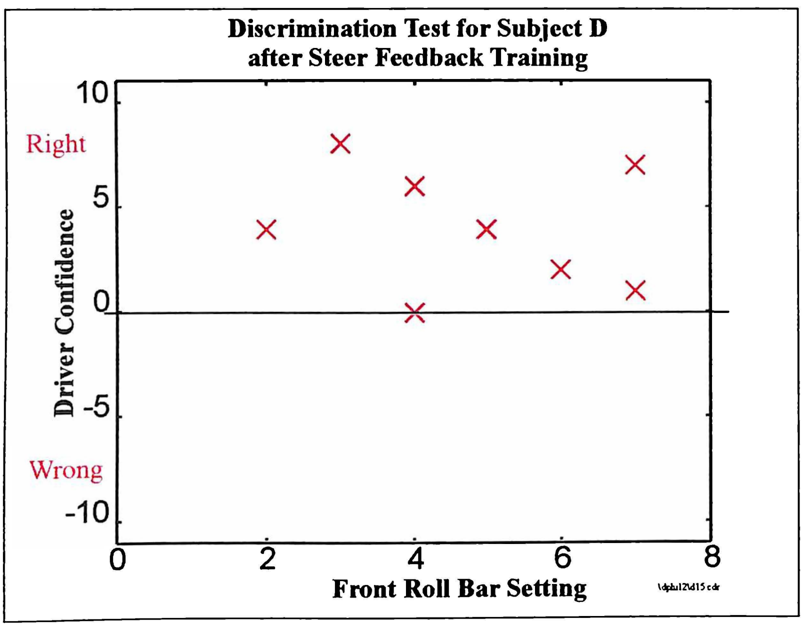


Figure 39

Figure 40 below shows the results of subject K's 2nd Discrimination test conducted at the reduced speed of 70 km/h.

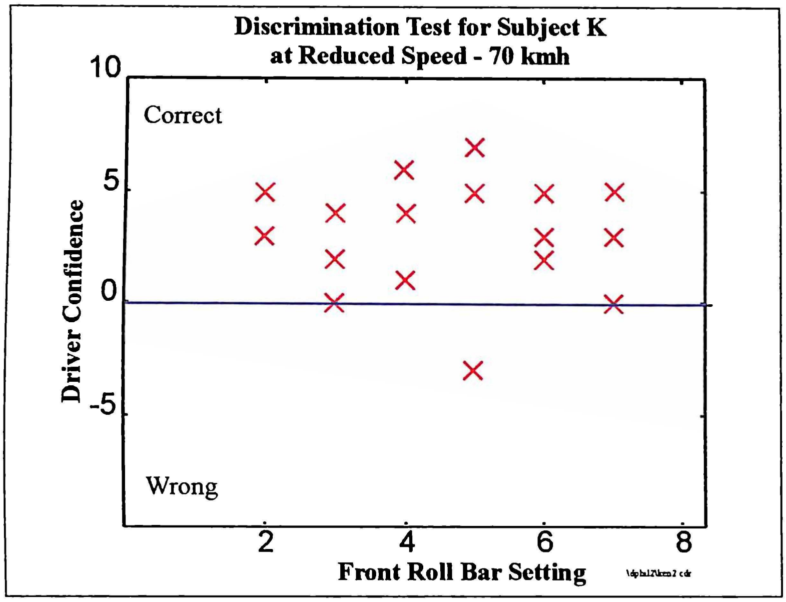


Figure 40

Figure 41 shows the Confidence- Sensitivity graph for subject S in her 2nd discrimination test when she concentrated on vehicle roll instead of vehicle understeer.

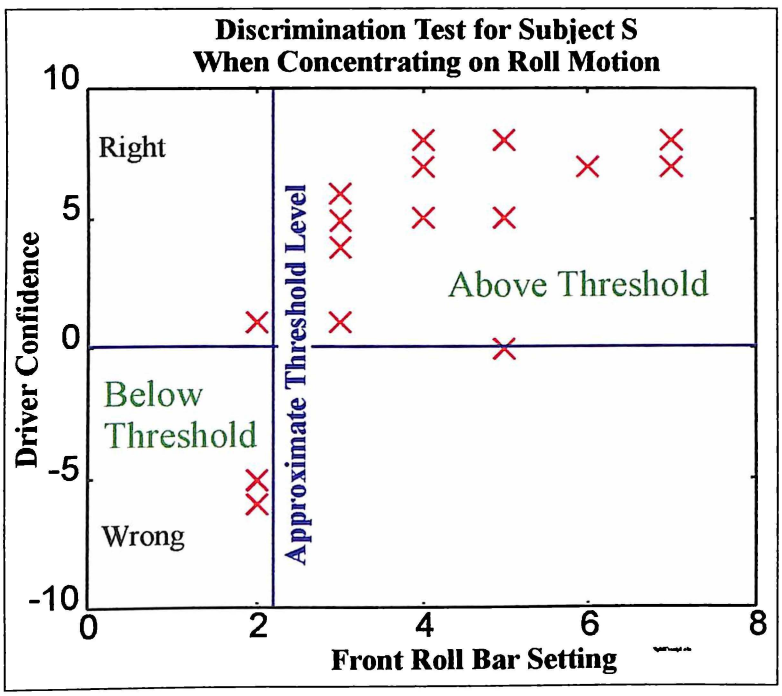


Figure 41

3.5 Discussion

3.5.1 Confidence vs Accuracy

The result of most concern for Ford Motor Co. was not that the VDEs scored at chance level, but that they did so with high confidence of being correct. VDEs' confidence ratings are the basis on which Ford allocates millions of dollars during the development process and from the results above it appears that they may be unrelated to the probability of being correct. Figure 32 shows that subject H was no more likely to be correct when he rated with a confidence of 10 than when he rated with a confidence of 2! Further, this overconfidence was extremely resistant to change. After each Discrimination session, subject H was shown his results. Despite 2 days of repeated failure he continued to report with high confidence. Typically what would happen is that he would think he had discovered what went wrong in the previous test and after the following session of self practice approached the next test with confidence. During the self practice subject H knew in advance what the Roll Bar settings were. It appeared to be the case that he found what he was looking for.

Prior to this study, no VDE conducted blind tests on vehicle modifications, apart from Jackie Stewart who is a consultant to Ford Motor Co. There were 2 reasons given for this lack of blind comparative testing;

i) That it is too time consuming and expensive.

ii) That the VDEs are experts and can reliably discriminate changes to a vehicle's set-up and hence it was not necessary.

As a consequence of this pilot study and the formal investigations detailed in later sections, Ford has now implemented blind testing during vehicle development.

3.5.2 Unaided Practice vs Training

Another finding which appears to be supported by the above data is that unaided practice may not be sufficient to achieve consistency of method when executing the Lane Change. Figures 33 and 34 show that there is a wide variation in both style and consistency between Jackie Stewart and subject H with respect to throttle control. Jackie was informally tested at Santa Pod during the calibration of the vehicle and modifications to the Lane Change layout. During that test Jackie correctly identified every Roll Bar setting with 100% accuracy when the most difficult Roll Bar procedure was used. Figure 34 shows that even during Jackie's very first set of runs he was consistent in his throttle control, particularly during the most important section of the Lane Change between 40 and 80 metres distance. This is in contrast to subject H whose throttle varies by a factor of almost 3 at 60 metres distance, despite 300 previous runs of practice. It would appear that this inconsistency was a standard part of subject H's normal evaluation procedure. Figure 35 shows that a short session of telemetry feedback training produced excellent throttle consistency for subject H.

This finding is repeated with the other professional test driver subject D. Figure 38 shows that subject D was almost twice as inconsistent on his steering inputs as subject S was during her first Discrimination test. Notice that subject D maintained this level of inconsistency for all 30 trials of the Discrimination test. Telemetry feedback training over 4 sessions of 10 trials concentrating purely on subject D's steering inputs produced an improvement in his steering consistency which was significant at the $p < 0.002$ level of significance (students t-test $N=80$, $t=14.9$). Figure 39 shows that after improving his steering consistency subject D scored 100% in his 2nd Discrimination test.

Subject B was given two short telemetry feedback training sessions consisting of only 10 trials each. In the first session he was told to concentrate on achieving a consistent throttle application while in the second he was told to concentrate on achieving consistent steering inputs. Subject B was given feedback after every second trial in both sessions. Figure 36 shows that during throttle training subject B improved his throttle consistency ($p < 0.01$) at the possible expense of some steering consistency (not statistically significant). During the steering training the original level of steering consistency was recovered without any degradation in the improved throttle control. This improved level of throttle consistency and the original level of steering consistency were maintained during the 2nd Discrimination test. Figure 37 shows that despite the improvement in throttle consistency, subject D was still not able to improve his discrimination performance.

Subjects H, D and B's results support the concept that telemetry feedback training might improve control consistency above the level achieved by unaided practice. It is worth remembering that subjects H and D are required to perform the Lane Change as part of their daily duties and have each performed this event considerably more than 5000 times each. It is unclear from the current pilot study whether or not these improvements to control consistency produce an improvement in discrimination performance, which is the key measure of a VDEs ability. That question was addressed in the subsequent main study.

3.5.3 Focus of Attention

Despite being a novice at the Lane Change, subject S displayed excellent control consistency right from the start of testing. This is shown for example in Figure 38 where her steering consistency is compared with that for the Lane Change expert, subject D. Despite this high level of consistency, subject S scored at chance level in her first Discrimination test session. During the first Discrimination test session, subject S based her answers on the apparent amount of understeer that the vehicle displayed, as did all the other subjects. Our inspection of the telemetry data indicated that the Roll Bar adjustments appeared to have a greater impact on the transient roll of the vehicle than it did on the understeer ratio. Instead of additional training to improve her control performance, subject S was trained to concentrate on the roll motion of the vehicle. With this different attention strategy, subject S improved her discrimination performance as shown in Figure 41. This result raised the possibility that subjects were attending to the wrong aspects of the dynamic motion and that attending to the correct aspects might improve their performance. This question was also addressed in the later formal study.

3.5.4 Driver Task Demands and Severity of Manoeuvre

Prior to this study, the accepted belief at Ford Motor Co. by senior VDEs was that vehicle handling tests should be conducted close to the performance limit of the vehicle (Ford Motor Co. personal communication). The argument promoted in support of this belief is that any deficiencies in the vehicle's design will be magnified at the vehicle's limit. By placing greater demands on the vehicle it was thought that any weaknesses would become more apparent. Taken to its extreme, this argument has some merit. If the vehicle is not exercised at all, i.e., remains stationary, then clearly the VDE cannot determine anything from the dynamic motion of the vehicle, as there is none. Similarly, at low speeds the vehicle behaves in a linear fashion and little can be learnt about vehicle instabilities and non-linearities. However, as the speed of the vehicle is increased, the demands made on the driver to control the vehicle also increase. It may be that at some very high speed the driver is allocating so much of their cognitive resource to controlling the unpredictable behaviour of the vehicle that there is insufficient resource left to evaluate the vehicle. Further, because the vehicle becomes more non-linear at higher speeds, it is also less predictable or repeatable for closely matched driver inputs. Mathematically we say the behaviour becomes more chaotic with increasing non-linearity. Thus it may be that beyond some speed the vehicle becomes more difficult to evaluate either because of the increasing task load placed on the driver or because of the increasingly poor correlation between the driver's inputs and the vehicle's dynamic behaviour.

This question was investigated with subject K. Figure 29 shows that at 75 km/h subject K scored at chance level while Figure 40 shows that when he was re-tested at 70 km/h he made only 1 error. To verify that this result was not due to some practice effect, subject K was again re-tested at 75 km/h. At this higher speed subject K again reverted to chance performance at all levels. It would appear at least for this inexperienced driver, that a lower manoeuvre speed in the Lane Change allowed him to more accurately assess changes to the Roll Bar.

3.6 Conclusion

The preliminary investigation indicated that;

i) The VDEs were not be performing as well as they were rating themselves

ii) It might be possible to improve VDE performance with specific training procedures over that attained by unaided practice

iii) Commonly accepted beliefs within Ford as to optimal testing procedures should be scientifically tested

iv) Altering the VDE's focus of attention might alter their discrimination performance.

v) Even if driver control consistency was necessary for good discrimination performance, it was not sufficient

4.0 Enhancements to Lane Change Discrimination Testing Procedures

4.1 Background

Analysis of the telemetry data from the investigation into VDE discrimination sensitivity and confidence ratings (section 3 above), showed that the Roll Bar adjustment used in that study to measure each driver's discrimination sensitivity was not ideal for a formal investigation of VDE performance or for evaluating the effectiveness of training procedures. In particular;

i) The adjustment range was not wide enough. Even using the largest difference between the Base and the Modified settings available, all drivers scored at chance level before consistency training because the difference between the Base and the Modified was too small. Unless we could accurately find every driver's threshold before and after each type of training program procedure, we could not evaluate the efficacy of that training intervention. Therefore we needed to find a new type of adjustment which produced a wider range of vehicle behaviour and which also had many more precise increments of adjustment.

ii) Changes to the dynamic behaviour of the vehicle caused by driving inconsistencies was too great compared to the changes caused by the vehicle adjustment. Small changes in the driver's style could easily mask the biggest change in the vehicle set-up. We wanted an adjustment which while affected by inconsistencies, was not completely dominated by them.

iii) There was too complex an interplay between the driver's inputs and the dynamic behaviours affected by the suspension adjustment. Some combinations of driver inputs made the vehicle behave as if it was in the Base condition when it was in the Modified condition and vice versa. If these combinations are not well understood then it is hard to explain to a driver why they got a particular answer wrong without a comprehensive analysis of all the telemetry channels for each run. A complex dynamical interplay requires a complex analysis before the training instructor can be sure that they are giving the subject the right advice. But one of the key factors in training a subject is to be able to quickly advise them after each pair of runs while the experience is still fresh in their minds. If it takes 10 minutes for the instructor to determine why the subject failed in a single pair of runs, then the training process becomes too slow for the subjects.

iv) Different driving styles caused the suspension adjustments to have markedly different effects on the vehicle. Some driving styles made the adjustments considerably easier to detect than other driving styles. This would make it difficult to compare perceptual performance across different drivers because they would be experiencing markedly different sensations. We needed an adjustment that was robust across different driving styles.

v) If we are going to extend our investigation into the efficacy of different training programs then the adjustment we make to the vehicle should be easily transferable across different vehicles. Fabricating different Roll Bars is a time consuming and costly exercise and will have completely different effects in one vehicle compared to another vehicle because of the differences in suspension geometry between different vehicles.

A considerable research effort was undertaken in England over a 3 month period at the conclusion to the preliminary VDE discrimination sensitivity investigation, to try to find a suspension adjustment that would satisfy the 5 requirements above. None of the suspension modifications tried during that research appeared satisfactory. However it appeared during other testing in New Zealand, that simply reducing the front tyre pressures produced a vehicle behaviour change that met the five requirements above. A study was conducted in New Zealand at the Meremere drag strip to verify whether this was the case. This study is detailed below.

4.2 Method

4.2.1 Subjects

5 Male (m) subjects and 2 Female (f) subjects were tested. They included;

2 Professional Race Driving Instructors (m)

1 Professional Flying Instructor (m)

1 Professional Race Driver (m)

2 Females and 1 Male whose occupations did not specifically require expertise in vehicle control.

The subjects were labelled d1 to d7.

4.2.2 Equipment

a) Vehicle and Instrumentation

A 1992 Ford Falcon XR-8 was used as a test vehicle. This vehicle had previously been modified by;

i) Fitting adjustable rate *Koni* Dampers front and rear

ii) Lowering the vehicle and reduction in suspension travel by approx. 2 inches

iii) The fitment of *Hankook* P125/60R15 93H steel belted radial tyres

The test vehicle was instrumented in the following manner;

i) A *Mitac* 386SX notebook computer with 40 MB HD was attached to the front passenger seat

ii) Telemetry data was fed to a *Pico* ADC-11 analog to digital converter which was attached to the *Mitac* notebook's parallel port

iii) Data Acquisition Software written by the author was used to sample the ADC -11 at 50 Hz and 10 bit resolution and to store the results on the hard disk

iv) A *Banner* 4- wire Multi-Beam Modular Photoelectric Retroreflective Scanner Block was mounted underneath the right rear of the test vehicle (in line with the middle of the driver' head) looking downwards at the road. This was set to send a trigger pulse to the Data Acquisition Unit (DAU) attached to the ADC-11 whenever the Falcon passed over a white reflective surface. Two white reflective surfaces were mounted on the Meremere Drag strip. One at the start of the Lane Change (to commence data acquisition) and one at the end of the Lane Change (to terminate data acquisition)

v) The following sensors were attached to the DAU;

  i) ADC Pin #3 = Channel 1 = Photoelectric Trigger

  ii) ADC Pin # 4 = Channel 2 = Longitudinal Acceleration

  iii) ADC Pin #5 = Channel 3 = Lateral Acceleration

  iv) ADC Pin #6 = Channel 4 = Longitudinal Velocity

  v) ADC Pin # 7 = Channel 5 = Lateral Velocity of rear of Vehicle

  vi) ADC Pin #8 = Channel 6 = Steering Position

  vii) ADC Pin # 9 = Channel 7 = Throttle Position

vi) Schaevitz +/- 5 g servo accelerometers were used. These were powered from the DAU so that 0 -1024 counts (full scale on the ADC card) corresponded to +/- 1 g Lateral or longitudinal Acceleration

vii) The Longitudinal and Lateral velocities were measured by the Correvit L and Q head sensors mounted on a specially fabricated bracket attached to the rear of the Falcon. A digital readout from the Longitudinal Velocity was attached to the right hand side of the front windscreen so the driver could monitor their entry and exit speeds without having to move his eyes away from the horizon

viii) A nylon split gear was manufactured and fitted around the steering shaft on split steel bearings which drove a second nylon gear attached to a high quality 270 degree ceramic potentiometer. This allowed +/- 135 of steering wheel movement to be measured by the ADC card

ix) The throttle position was determined by taking a voltage measurement from the Electronic Control Unit of the Fuel Injection System

b) Test Track

The layout of the Lane Change was the same as that used during the preliminary investigation into VDE discrimination sensitivity in Florida, which is shown above in Figure 25.

4.2.3 Procedure

The testing procedure was divided up into two phases;

i) Consistency practice session

ii) Discrimination test session

a) Consistency Practice Session

The Lane Change manoeuvre is a very severe event simulating an accident avoidance manoeuvre and therefore most subjects require approximately 10 trials before they can complete it at the correct speed. Each subject was given 26 practice runs during which they were told to gradually increase their speed until they attained the target entry and exit speed. Because the Ford XR-8 is a sports performance car while the Ford Explorer is a light truck, the speeds at which the Lane Change event should be completed in the XR-8 are higher than the speeds for the Explorer. The subjects were told that during the Discrimination test session they would have to ensure that their entry and exit speeds were as close to 90 km/h as possible.

The subjects were advised that the purpose of the Consistency practice session was;

i) To enable them to become familiar with the handling characteristics of the XR-8

ii) To attain as consistent a line as possible through the Lane Change

The subjects were advised that after the 26 practice runs in the training session they would be tested to find their sensitivity to changes in the front tyre pressures of the XR-8 (Discrimination test session). It was stressed to the subjects that in order to discriminate between the tyre pressure changes they would have to be able to drive each run in a very similar manner otherwise changes in their driving style might mask the changes induced by altering the tyre pressures. Throughout the Consistency practice session, all the tyres were set at 36 psi, which corresponded to the Base condition.

b) Discrimination Test Session

The Discrimination test session required each subject to perform pairs of runs. One of these runs would be with the vehicle in the **Base** condition (36 psi) as per the Consistency practice session and one of the runs would be with the vehicle in a **Modified** condition which was achieved by lowering the front tyre pressures. Subjects were required to identify which of the 2 runs was the Base run using the 2 alternate forced choice procedure. Tyre pressures were altered on the next pair of runs using the staircase procedure (Cornsweet, 1962). Trials continued until there were 3 successive reversals or the subjects reached the maximum Modified pressure of 32 psi. In addition to determining which was the Modified and which was the Base run, the subjects had to rate how confident they were of their choice using the same 0 - 10 scale used in Florida. (section 3.2.3).

4.3 Analytical Techniques

In the previous study into VDE discrimination performance (section 3), the VDEs reported that the most noticeable effect of adjusting the Roll Bars was to change the vehicle's steering responsiveness during the middle or transition section of the Lane Change. If a vehicle has less steering responsiveness, the driver can compensate for this by using a combination of 2 strategies;

i) Using a larger Steer Angle

ii) Holding a similar Steer Angle on for a longer period

For example Figure 42 shows the steering traces for two runs from the above study.

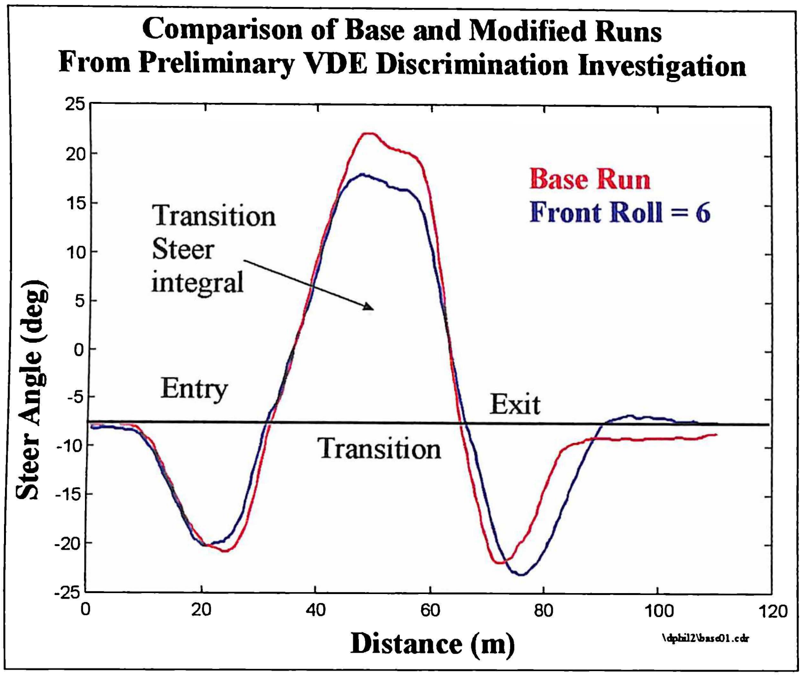


Figure 42

In Figure 42 we see that in the Base run the driver has used a larger Steer Angle during the transition section (which corresponds to region marked section 3 in Figure 25) than they did in the Modified run. This increase in Steer Angle for the Base run will mean that the area under the Steer Angle -Distance curve during the transition section, called the Transition Steer integral, will be larger for the Base run than for the Modified run. In this case we say that the vehicle had less steering responsiveness in the Base condition than it did in the Modified condition.

Similarly, if a driver use the same amount of Steer Angle but holds it on for a longer distance in the transition section because the vehicle does not respond as well, then the Transition Steer integral will also increase. This second driver strategy to cope with reduced steering responsiveness is shown below in figure 43.

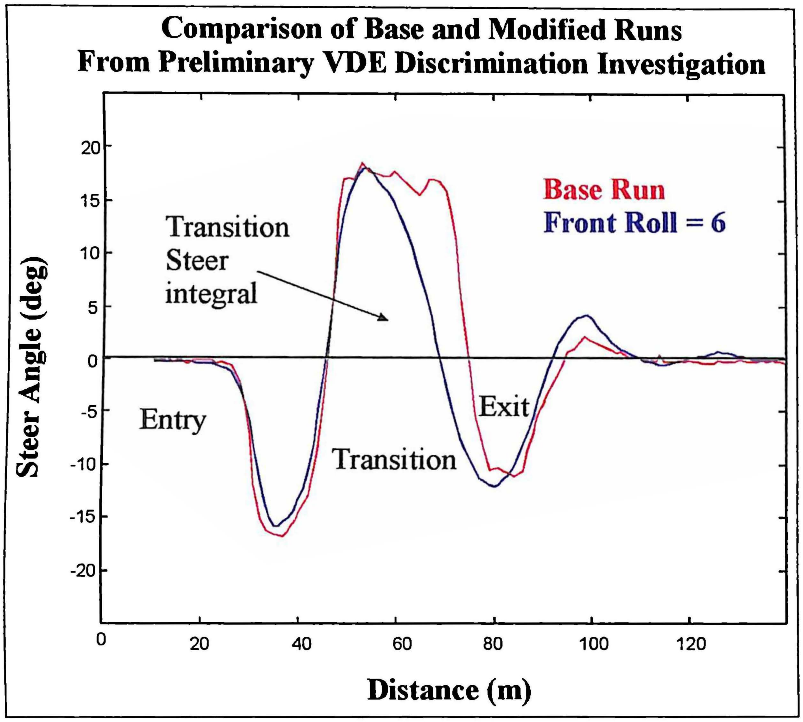


Figure 43

Thus the Transition Steer integral is a good measure of the vehicle's steering responsiveness that is independent of the strategy that the driver used to negotiate the cones. This concept of the Transition Steer integral was developed by myself during 1993 and has subsequently proven to be useful for a number of procedures and now forms part of Ford's standard vehicle evaluation metrics.

Figure 44 below shows how the Transition Steer integral was affected by different front Roll Bar settings during the preliminary investigation (section 3 above). In Figure 44 we have averaged the Transition Steer integrals for all drivers at each Roll Bar setting.

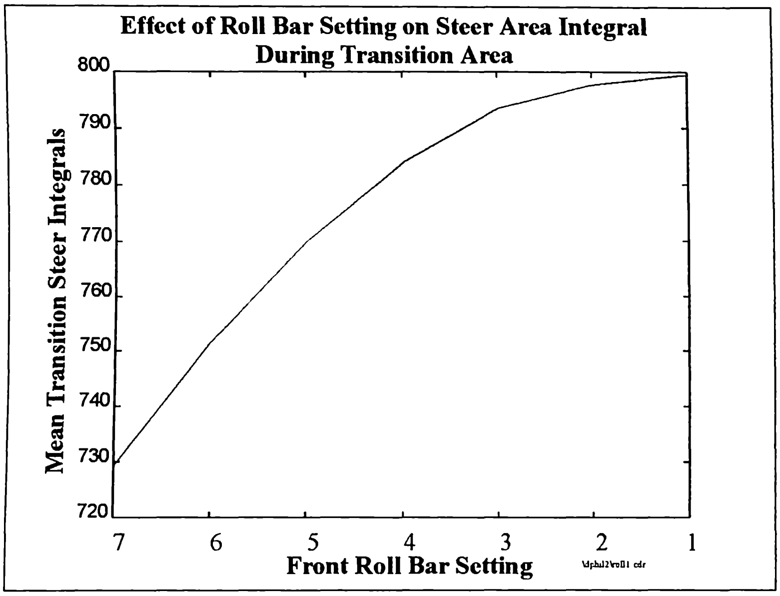


Figure 44

Figure 44 shows that stiffening the front Roll Bar from setting 1 (Base) to setting 7 reduced the Transition Steer integral by approximately 9%. Thus, increasing the Roll Bar stiffness from the Base setting to a harder Modified setting increased the steer responsiveness of the Explorer.

4.4 Results

The new methodology allowed each subject's discrimination threshold to be reliably determined regardless of their skill level. For example Figure 45 shows the Confidence-Sensitivity graph for subject d5, the subject with the lowest performance, while Figure 46 shows the corresponding graph for best subject d2 who was a Professional racing driver. Thus there was sufficient range in the vehicle adjustment to allow us to measure all subjects. We could of course have lowered the tyre pressures even further if a subject with worse performance was encountered.

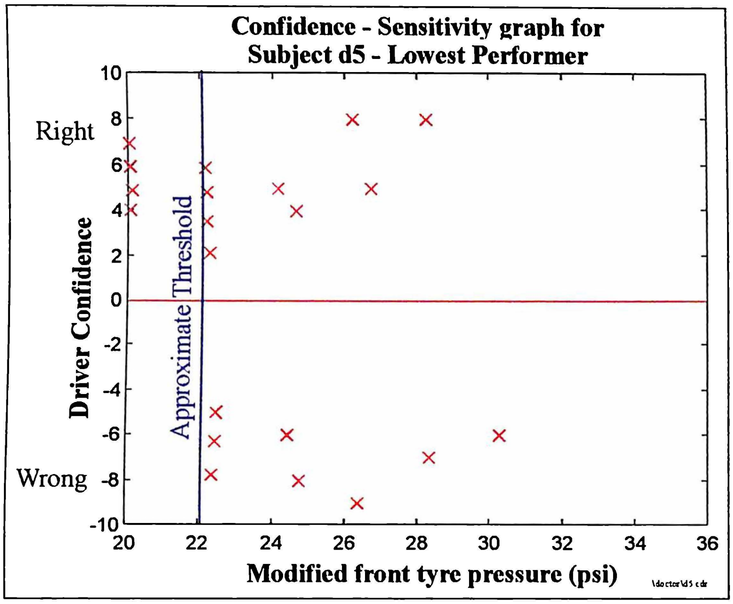


Figure 45

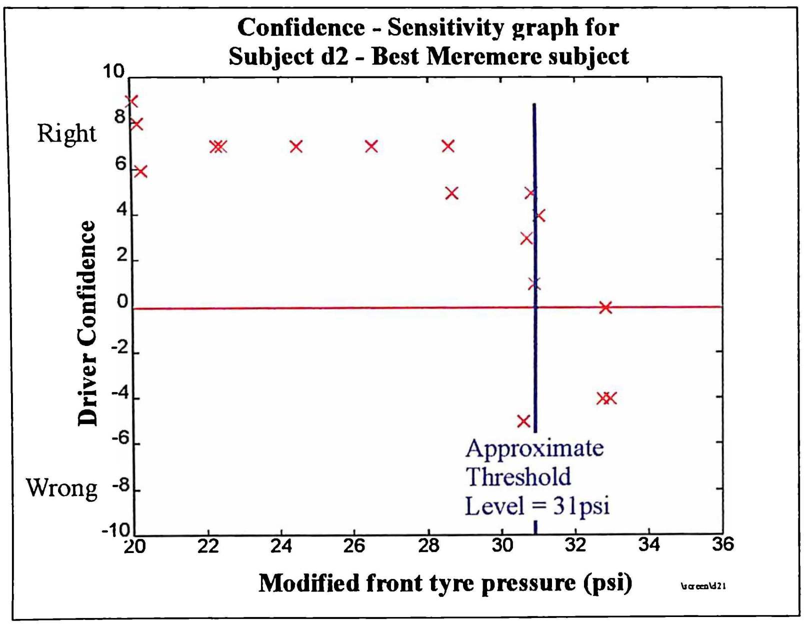


Figure 46

Figure 47 shows the linear best fit regression lines of Transition Steer integral versus Modified tyre pressure for each of the 7 drivers.

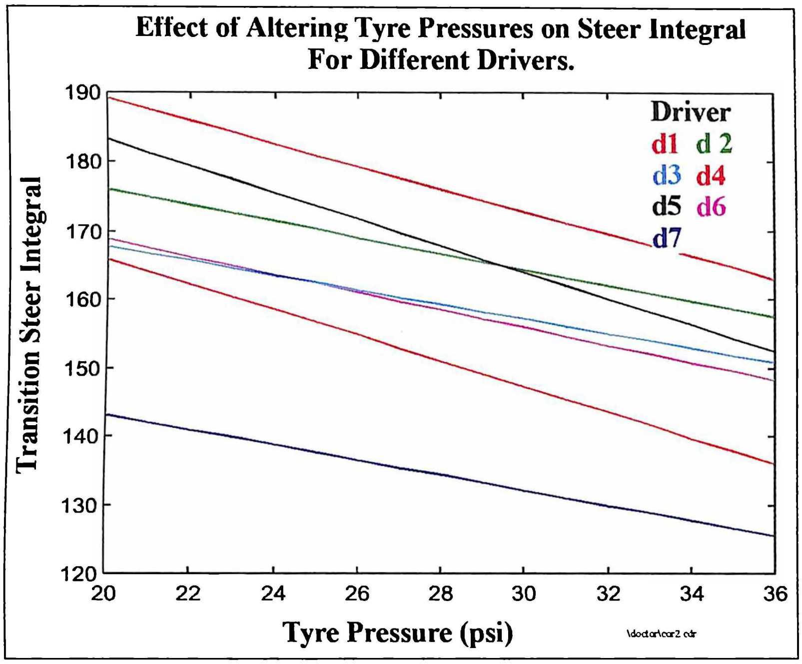


Figure 47

The corresponding correlation coefficients for each driver for these regression curves was;

Driver 1	r = - 0.8896
Driver 2	r = -0.8211
Driver 3	r = -0.8346
Driver 4	r = -0.6978
Driver 5	r = -0.8629
Driver 6	r = -0.9284
Driver 7	r = -0.7771

Thus lowering the front tyre pressure from the 36 psi Base setting reduced the steering responsiveness of the test vehicle which caused a corresponding increase in the Transition Steer integral.

Figure 48 shows the linear best fit regression lines relating entry speed to Transition Steer integral for each of the 7 drivers.

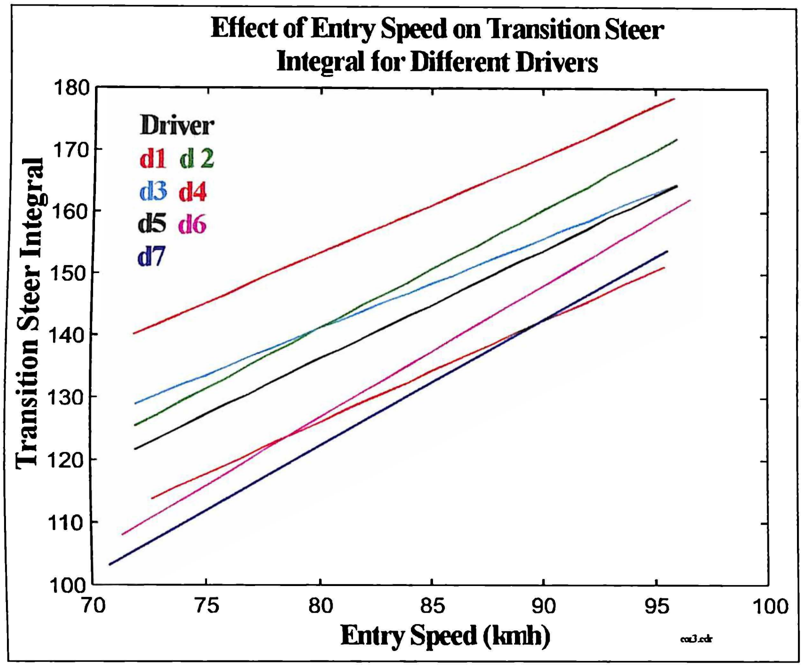


Figure 48

Data for Figure 48 was obtained from both the Consistency practice session where the drivers were learning to negotiate the Lane Change at a lower speed and from Discrimination test session where the drivers were required to keep as close to 90kmh as possible. Figure 48 shows that the Transition Steer integral increases as the speed of the manoeuvre is increased because the higher speed places a greater demand on the vehicle. The most important aspect of Figure 48 is the similarity of the gradients.

Figure 49 shows the effect of Peak Entry Steer Angle on the Transition Steer integral for all 7 drivers. Each driver's runs include both Base and Modified tyre pressure runs.

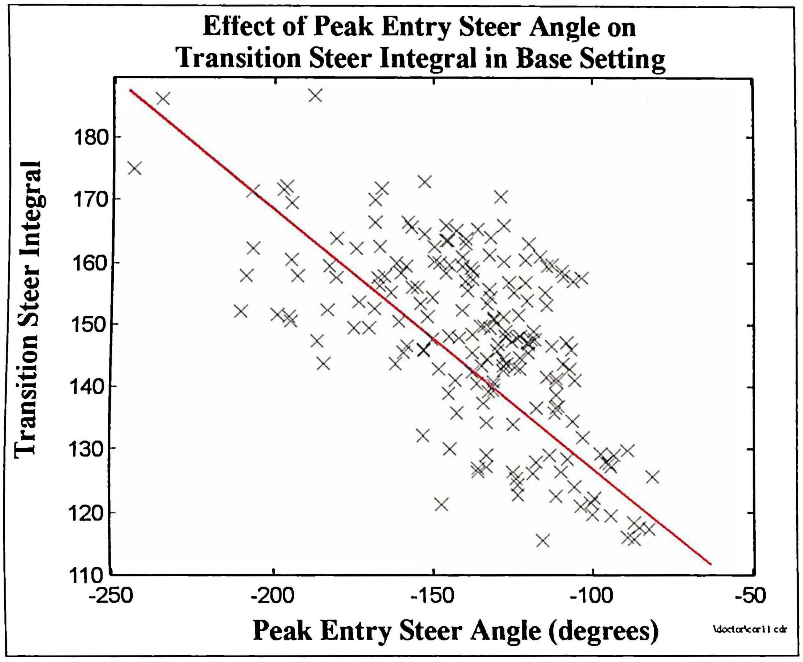


Figure 49

Figure 49 shows that when a driver increases the Peak Entry Steer Angle by turning in harder, this places a greater demand on the vehicle which in turn increases the Transition Steer integral. Because Figure 49 was derived from all 7 drivers and shows that the Transition Steer integral it is also dependent on the driver's style as well as on the tyre pressure. If the driver uses a greater Peak Entry Steer Angle then the Transition Steer integral will inevitably be increased.

4.5 Discussion

Figures 47 and 48 allow a comparison of the relative sensitivity of the vehicle's behaviour to tyre pressure adjustments and to driving inconsistencies. Comparing the gradients in Figures 47 and 48, we see that a speed change of approximately 10 km/h produces the same change in Transition Steer integral as altering the tyre pressures from the softest to the hardest settings (20psi to 36 psi). These 2 graphs can be compared to Figure 50 below which was obtained from the preliminary investigation into VDE discrimination sensitivity conducted in Florida (section 3).

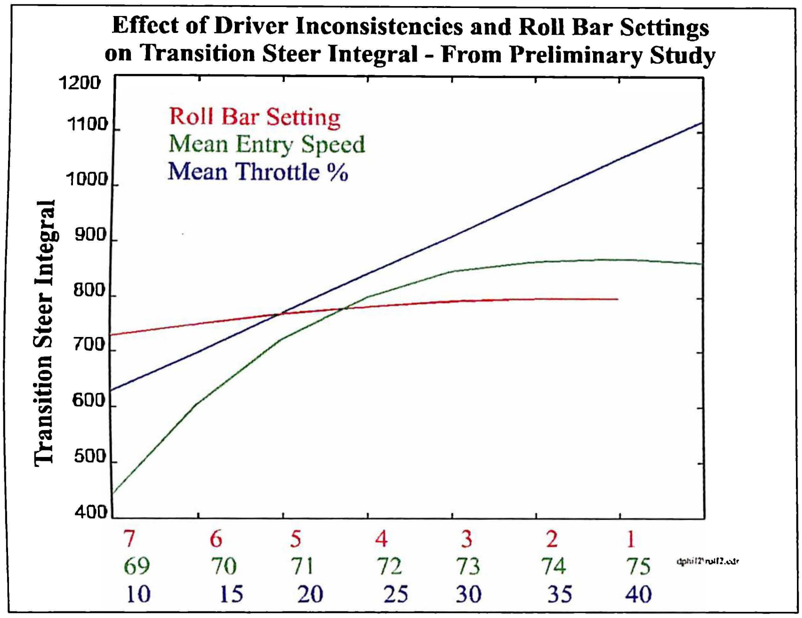


Figure 50

Figure 50 shows that a 3 km/h variation in driver entry speed in the preliminary investigation (section 3) produced a greater change in Transition Steer integral than that produced by changing the Roll Bars from the softest to the hardest Roll Bar settings. The red line in Figure 50 is identical to the red line contained in Figure 44 , but the shape of the curve in Figure 50 is difficult to detect because of the large scale needed in Figure 50 to show the greater sensitivity of the vehicle to driver inconsistencies (speed or throttle variations).

Thus, the new methodology of altering front tyre pressures appears to provide a vehicle adjustment that was sensitive to driver inconsistencies without being too sensitive as as occurred during the preliminary investigation. In the preliminary investigation a very small inconsistency in a driver's style would completely swamp any effect caused by altering the Roll Bar settings. Thus our first objective for enhancing the test procedure was met.

Figures 47 and 48 also show that there were large differences in driving style between the different subjects as would be expected with drivers whose backgrounds ranged from a Professional racing driver to subjects who had never driven a vehicle on a track before. For example Figure 47 shows that the Transition Steer integral ranged from a mean of 142 for one driver to a mean of 189 for another driver at 20psi. Despite these large differences in driving style ($p < 0.001$), altering the tyre pressures for each driver produced similar percentage changes in vehicle behaviour. This is exemplified by the gradients in Figure 48 which show that an increase in entry speed produced similar percentage increases in Transition Steer integrals for all the drivers. Thus, the primary metric which the drivers were sensitive to was approximately equally affected by different driving styles. One driving style did not confer a dramatic advantage over another in terms of discrimination masking.

Finally Figure 49 shows that despite the very wide variation in driving styles (peak entry steer angle of -80 degrees to peak entry steer angle of -250 degrees) that there is a good correlation of peak entry steer angle with Transition Steer integral. These well behaved and easily understood correlations make it easy for the instructor to identify how a particular subject's driving style is influencing the behaviour of the vehicle and hence the sensations that they are experiencing.

Note: The units of Transition Steer integral used in the two studies (Meremere and Florida) were different because in the current Meremere study the product of Steer Angle x Distance moved was converted to degree- metres whereas in the preliminary investigation in Florida the product of Steer Angle x Distance moved was reported in raw telemetry units of steering angle (telemetry encoding counts) x distance moved (m). Further, the sampling rate in the preliminary study was 20 Hz while in the Meremere study it was 50 Hz. The actual value of the steer integrals will also differ because of the different behaviour of the XR-8 compared to the Explorer. I have preserved the old Ford raw units of Transition Steer integral from the 1993 preliminary study for 2 reasons;

i) The data is confidential to Ford which makes it difficult and expensive to obtain all the data for conversion to the Meremere units

ii) To maintain compatibility with previous reports produced for Ford on the preliminary study

However, while the use of different units is not desirable it is not a problem for our discussion as we are not comparing one vehicle with another, which would require consistent scaling. Rather what we are interested in is the <u>percentage</u> change in the values resulting from either the driver's style (speed, steering input or throttle input) or changes to the vehicle (Roll Bar setting or tyre pressure). The percentage change is of course unaffected by the scaling used.

4.6 Conclusion

The new methodology of adjusting tyre pressures appears suitable for evaluating subjects' discrimination performance and the efficacy of different training procedures.

Section B

Main

VDE

Investigation

5 Outline of Main Study

24 Professional VDEs were selected by Ford Motor Co to attend a training and research program conducted by the author in England during May - November 1995. Four separate classes each lasting one week and containing six subjects were conducted.

The purpose of the program was;

i) To improve the VDEs performance in evaluating a vehicle (training)

ii) To determine the psychological factors that affected or predicted VDE performance (research)

iii) To determine the effectiveness of different training programs on improving vehicle evaluation performance (research)

Subjects were administered both laboratory and practical driving tests.

The laboratory tests consisted of;

i) Vestibular supra-liminal discrimination tests which were conducted at the RAF Aeromedical Facilities at Farnborough

ii) The computerised Basic Attributes Tests that are used to select fighter pilots which were conducted at the RAF Directorate of Recruiting and Selection at Cranwell.

The practical driving evaluation consisted of the Lane Change Discrimination test to measure the subject's vehicle evaluation performance and was conducted at the RAF Upper Heyford Airfield. Each subject's Lane Change performance was measured both before and after training.

6 Vestibular Experiments

6.1 Introduction

6.1.1 The Importance of the Vestibular Sense

The original reason for performing the vestibular tests was to determine whether a subject's supra liminal vestibular discrimination performance was related to their ability to evaluate vehicle changes. If a relation was found, then vestibular tests could be used to help select the most suitable subjects for future VDE training. A review of the literature showed that there were good reasons for hypothesising that vestibular performance might be related to a VDE's vehicle evaluation performance.

Walsh (1961) showed that the non-visual threshold for perception of linear motion by subjects who also had non-functional vestibular systems was at least 10 times larger than in normal subjects for frequencies below 1 Hz. For rotational motion about the z-axis, this threshold was raised by a factor of over 100. Therefore it would appear that the vestibular system might play an important role in motion perception above liminal values in the absence of visual information. Furthermore, the somatosensory mechanoreceptors do not appear to contribute to the liminal perception of movement at frequencies below 1 Hz because the thresholds of patients with high spinal lesions were found to be similar to those of normal subjects (Walsh 1961).

But the question arises as to whether the vestibular system contributes significantly to any perceptual or control tasks where vision is fully available, particularly in a situation similar to that where a VDE is controlling a vehicle.

An extensive series of studies was conducted at Delft University over a period of 4 years on roll tracking and roll perception under combinations of vestibular, central visual information and peripheral visual information (Hosman and Van Der Vaart 1981, 1983, 1984). In these experiments, the subjects sat inside a motion simulator and were provided information from three sources;

i) A central display containing an artificial horizon

ii) A peripheral visual display from the outside world

iii) Cockpit roll motion as perceived by the vestibular system

The subjects were tested under two conditions;

i) A nulling task where they were required to keep the simulator cockpit vertical in the presence of a random roll disturbance

ii) A tracking task, where the subjects had to make the cockpit follow a random position indicated on the central artificial horizon display

For the nulling task, the subjects were tested in all 7 combinations of available information. Figure 51 below shows the standard deviation of the tracking error under each of the 7 information conditions (vertical bars are the standard deviations in these tracking errors across subjects).

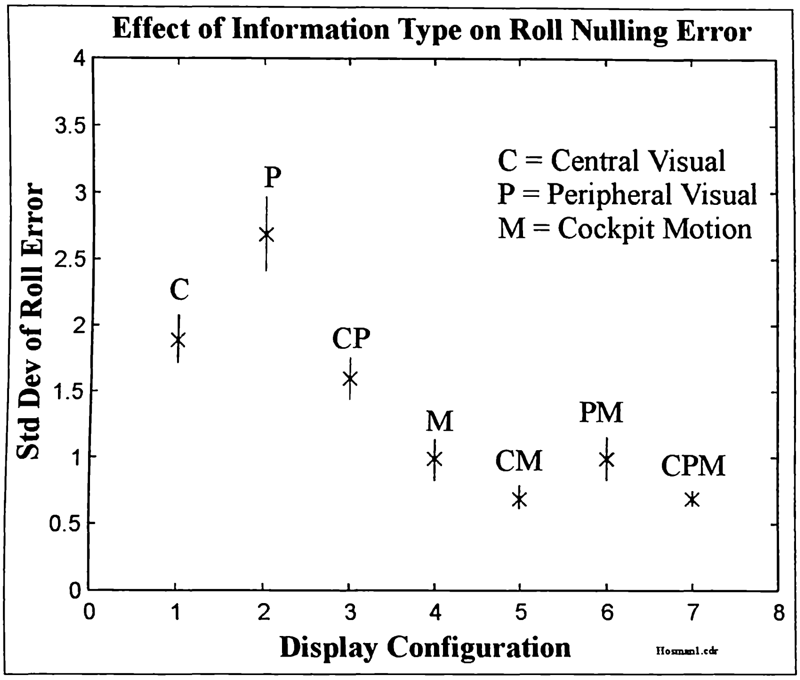


Figure 51 (Hosman and Van der Vaart, 1990)

Figure 51 shows that motion information had a significant (at $p < 0.05$) impact on a subject's ability to null out the perturbing influence. In particular, the mean roll error when the subject was presented with only vestibular motion information was significantly less than the mean roll error when either central or peripheral or both central and peripheral visual information alone were available to the subject.

In the tracking task of course, the central display had to be present in all cases, as this was the stimulus that the subjects had to follow. Nevertheless, Figure 52 below also shows an effect of providing motion information to the subject in performing this task.

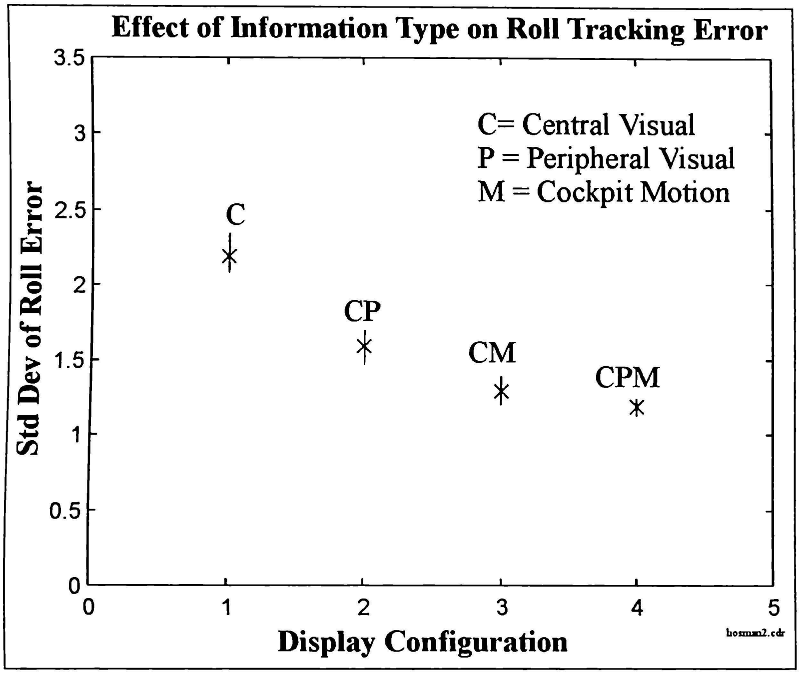


Fig 52 (Hosman and Van der Vaart, 1990)

The behaviour of the human operator under all these conditions can be modelled by the quasi human operator model developed by McRuer. (McRuer et al., 1965, and Mc Ruer and Krendel, 1974). In this model, the human operator response, $H_p(s)$ is given by;

$$H_p(s)=K_p(1+T_L s)/(1+T_i s)\ e^{-ts}$$

where $T_i s$ = Human Operator Lag Time

$T_L s$ = Human operator Lead Time

t = effective time delay which accounts for all delays in perception, processing and control output generation.

K_p = static controller gain

Using this model it is possible to calculate the subject's cross over frequency (w) and the phase margin (p) as well as the parameters describing the operator function.

Table 1 tabulates these parameters for each of the various information conditions in disturbance nulling task from Figure 51 while Table 2 tabulates the parameters for the target following task corresponding to Figure 52.

Display	w	p	K	TL	Ti	t	Std Error
C	3.18	14	1.3	0.7	<0.1	0.28	1.94
P	3.16	15	1.3	0.7		0.28	2.67
CP	3.55	15	1.3	0.7		0.25	1.64
M	4.64	21	2.7	0.5		0.15	0.99
CM	4.76	19	2.6	0.5		0.18	0.79
PM	4.89	19	2.6	0.5		0.15	1
CPM	5.03	18	2.6	0.5		0.16	0.7

Table 1 Quasi-Linear Pilot Model Parameters for Nulling Task

(Mc Ruer and Krendel, 1974)

Display	w	p	K	TL	Ti	t	Std Error
C	2.26	15	0.55	1	< 0.1	0.35	2.23
CP	2.24	35	0.15	4		0.35	1.63
CM	2.23	39	0.09	6		0.35	1.32
CPM	1.66	56	0.07	6		0.35	1.26

Table 2 Quasi-Linear Pilot Model Parameters for Tracking Task

(Mc Ruer and Krendel, 1974)

Table 1 shows that the addition of Peripheral visual information to the Central visual information lowers the effective time delay by 0.03 seconds while the addition of the Vestibular signals lowers the effective time delay by a much larger factor of 0.10. This reduction in effective time delay produced by the addition of the vestibular signal, while it appears quite modest, does allow the subject to significantly increase their static gain which in turns increases their cross over frequency. A high cross over frequency is an indication of a high controller gain over a large frequency range which is desirable for accurate error compensation (Sheridan and Ferell, 1974). This indicates that the vestibular system has an important role in motion perception and vehicle control over and above its ability to simply detect motion, which the visual system can perform remarkably well on its own. What the above experiment suggests is that the vestibular system provides high frequency cues to supplement the accurate but low frequency visual cues. This combination provides a wide band sensory system giving accurate response over a remarkably diverse range of motion inputs or target goals.

Similar research was conducted by Zacharias and Young (1981) who placed subjects in a LINK GAT-1 small aircraft trainer. A stripe pattern was projected onto the windows of the aircraft which could be made to rotate relative to the windows independently of the Yaw motion of the trainer platform. The subjects were given the task of nulling the applied randomly perturbed visual pattern either with or without related vestibular cues. That is the trainer could be held constant and the visual stripes moved about the trainer to provide a visual only stimulus, or the trainer could be moved and the motion of the visual stripes adjusted in the opposite direction to give the illusion of moving against a visual background which was fixed in space. Zacharias and Young performed FFT's on the outputs under each condition to compute the operator transfer functions under the assumption that the nulling task is the sum of both the vestibular and visual cues and hence calculated the input describing functions. This analysis also showed that the visual system had a high gain at frequencies below 0.02 Hz, but that above that frequency the vestibular system had significantly higher gains than the visual system.

Meiry (1965) performed a related experiment where subjects were required to control the motion of a cockpit simulator. A second order tilt instability was introduced of the form $2w^2/(s^2-w^2)$ where w is the frequency of divergence. Undamped divergent frequencies between 7 and 23 cpm were used. In the vestibular only condition, the subject sat inside the cockpit blindfolded and had to keep the cockpit vertical by use of a joystick. In the visual and vestibular condition, the subject could look out of the cockpit window to the laboratory wall 10ft in front of him which had horizontal and vertical reference lines painted on it. In the vision only condition, the subject sat 10ft behind the cockpit which had a vertical reference line painted on it. Figure 53 below shows the root mean square tracking error under the 3 different conditions.

With vision alone, the subjects were unable to control the motion of the cockpit when the frequencies of divergence exceeded 17 cycles per minute. With vestibular information alone the subjects could control the cockpit at all frequencies. Addition of visual information to the vestibular information produced only a small increase in performance.

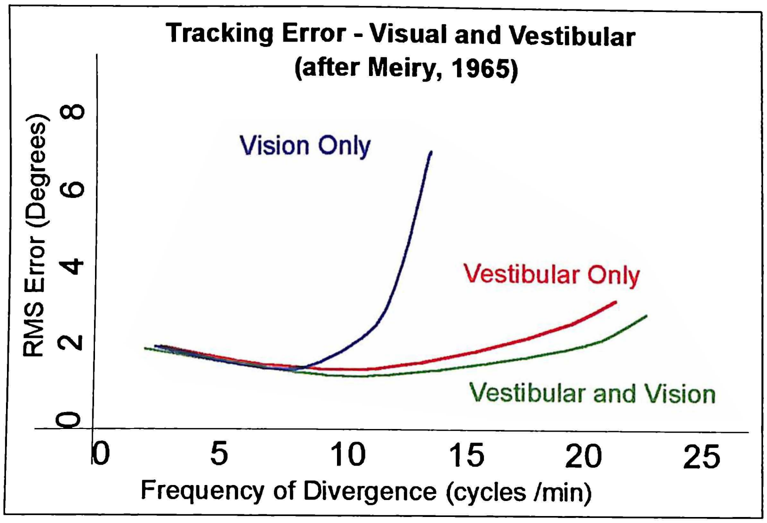


Figure 53

Thus the vestibular system appears to have a significant role in a human's perception and control of a vehicle particularly in terms of high frequency corrective responses and as such should be a good candidate for investigation for prospective or current VDEs.

However the vestibular performance attributes that we wished to measure for each subject were different to those evaluated by the standard vestibular tests administered by clinical otolaryngologists or those published in aerospace journals for astronauts or pilots. We were not interested in a subject's absolute threshold for detection of minute motion, nor tolerance to extreme motion, but rather we were interested in a subject's ability to distinguish between two similar but slightly different motions that were significantly above their detection threshold. When a test driver reports that one vehicle design behaves in a different fashion to another, the language that they use in describing these differences, is often of small incremental changes of motion within a large easily detected (and sometimes violent) motion. For example, when comparing two vehicle set-ups, a subject may report that the Yaw Rate response of the vehicle at the exit section was lower in one condition than in the other condition. In this case, the Yaw Rate is easily detected in both runs, but the driver has judged that the Yaw magnitude at one specific section of the Lane Change is greater in one run than in the other run. It cannot be assumed that a subject's performance on the simple detection of the presence or absence of motion at liminal levels will be related to their ability to discriminate between the magnitude of two motions at high stimulus levels. Therefore the subjects' **Amplitude Discrimination** performance was measured.

After reviewing the performance of Jackie Stewart in the Lane Change during our preparation for the preliminary investigation, another consideration emerged when we designed the vestibular tests. Not only did Jackie appear very sensitive to changes in motion, but his verbal descriptions of events contained far more detail than anyone else. He would describe little bumps and wiggles in the telemetry plots as each of the various suspension components contributed to the motion of the vehicle during the Lane Change manoeuvre. That is, not only did he appear to be able to tell the difference between the magnitude of two stimuli, he also appeared able to detect, encode and store much more of the fine structure contained within the motion stimulus. His encoding scheme and motion memory allowed him to describe the motion in far greater detail. Therefore, we designed another type of vestibular discrimination test which we called the **Embedded Discrimination** test to measure each subject's ability to identify small perturbations embedded inside a larger constant stimulus. In this test, the same large stimulus used for the Amplitude Discrimination test was presented, but in addition a smaller perturbing stimulus was superposed within one part of the wave form to alter its shape. The subject's task was to identify which of two presentations had the largest perturbation embedded inside the main stimulus.

Because the detection of rotation is primarily the responsibility of the semi-circular canals while the detection of lateral motion is the responsibility of the saccule and utricle, the possibility existed that a subject could have good performance in a rotary test of magnitude discrimination and poor performance in a lateral test of magnitude discrimination. Therefore, it was decided to test subjects' performance on both the Rotary motion platform and the Linear motion platform.

There were thus 4 different test procedures;

i) Linear Platform - Amplitude Discrimination

ii) Linear Platform - Embedded Discrimination

iii) Rotary Platform - Amplitude Discrimination

iv) Rotary Platform - Embedded Discrimination

A brief review of the physiology and absolute detection performance of the rotary and linear components of the vestibular system follows by way of introduction to our study.

Figure 54 below shows a schematic representation of the vestibular system.

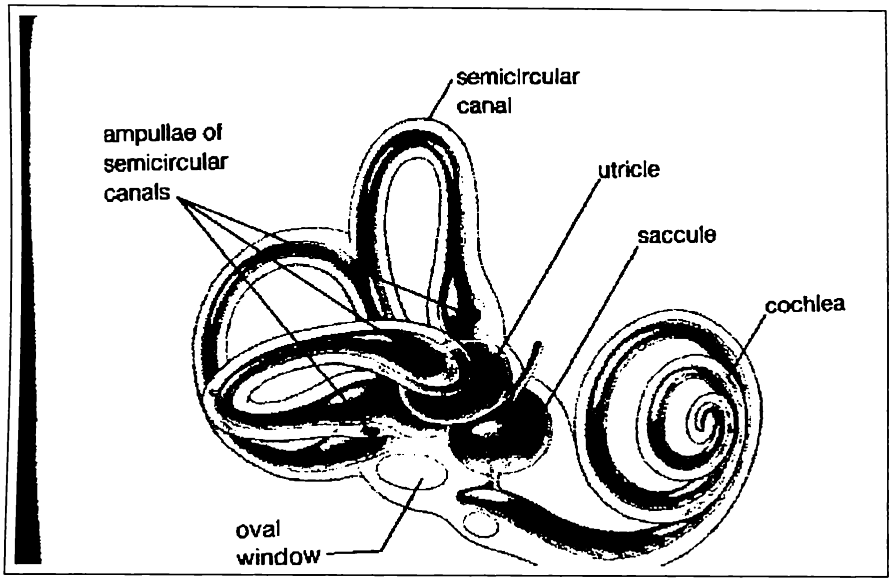


Fig 54 Schematic Representation of the Vestibular System
(Taken from Graham, 1990)

The 3 semi-circular canals detect rotary motion in each of the three axes, Yaw, pitch and roll while the utricle and saccule detect lateral acceleration in the x,y and z axes.

6.1.1 Vestibular Rotary Motion Detection

Figure 55 below shows a schematic of the semicircular canal which shows how rotary motion is detected. The semi-circular canal is filled with a fluid called the endolymph. At the base of the semi-circular canal, there is a swelling called the ampulla which contains the receptor organ called the crista. Mounted on top of the crista is the cupula which is a gelatinous body that obstructs the motion of fluid in the semi-circular canals. When the vestibular apparatus is rotated about the axis of one of the semi-circular canals, the inertia of the fluid contained in that canal will cause it to flow past the cupula making the cupula deflect to one side.

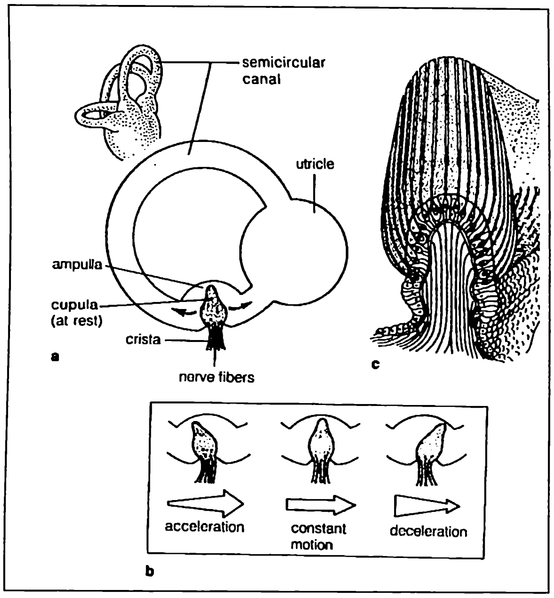


Fig 55 Schematic of Semi-circular canal
(Taken from Shiffman, 1976)

The deflection of the cupula distorts the crista which contains specially designed sensory cells. Each of these sensory cells have approximately 60 - 100 cilia rising from their surface which have a directional sensitivity (Lindeman, 1969). Displacement of the cilia to one side increases the resting discharge rate of the sensory cell causing it to fire while displacement in the opposite direction decreases the rate of firing. This discharge rate is a monotonic function of the mechanical deformation. Interestingly, while *angular acceleration* is the stimulus for exciting the semi-circular canals, the rate of firing of the sensory cells in the crista tends to be proportional to *angular velocity* rather than acceleration. The reason for this is that the hydrodynamics of the duct-cupula-endolymph system are very heavily damped (Benson, 1990). Because the system is so heavily damped, the fluid does not immediately attain a rotational velocity past the cupula equal to that of the rotation of the vestibular system itself. Rather, the endolymph gradually accelerates so that the velocity past the cupula and hence the deflection of the cupula depends on both the rate of acceleration (α) of the canals and also the time (τ) for which the system has been undergoing this angular acceleration. As $W=\alpha\tau$ then the system will appear to respond to angular velocity (w) if the angular velocity is maintained for a moderate time period (Huang and Young, 1981,; Meiry, 1965; Clark and Stewart, 1974 and Fenessy, 1975). Of course, if the angular acceleration ceases and the system maintains a constant angular velocity for a lengthy period of time, then the endolymph will gradually catch up with the rotation of the system and will then no longer deflect the cupula. The time period for this is of the order of 10 seconds, depending of course on the rate of acceleration and hence this habituation does not occur normally in subjects during self generated motion (Benson, 1990).

The absolute threshold for detection of rotary motion is very dependent on the experimental technique used. Factors such as whether visual information is available (Zacharias and Young, 1981) and the nature of the motion stimulus, (whether oscillatory or cosine bell etc.) have a profound effect (Clark, 1970). If the stimulus duration is less than 5 seconds, then it has been found that the detection threshold is a function of the peak angular velocity reached during the stimulus presentation and further that this threshold decreases as the frequency of the stimulus is increased. Figure 56 shows the mean angular velocity (deg/sec) threshold for detecting a single cycle of a sine wave recorded for 6 subjects.

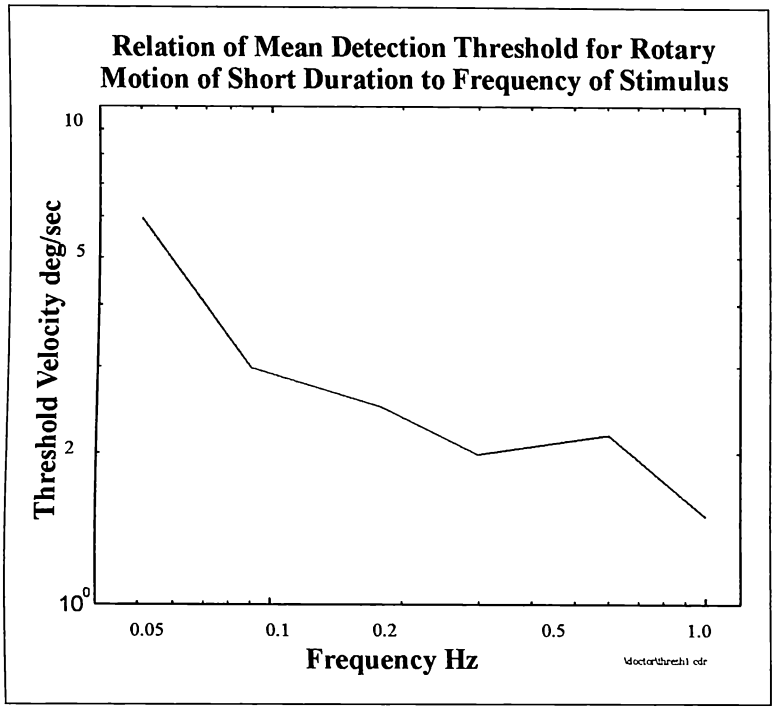


Fig 56 (from Benson, Hutt and Brown, 1989)

If the detection threshold was dependent solely on peak angular velocity, then fig 56 would have a gradient of zero. Conversely, if the detection threshold was dependent solely on the peak angular acceleration, then the gradient of Figure 56 would be -1 log unit / decade. In fact the gradient of Figure 56 is approximately -0.2 log units /decade between the frequencies of 0.1 and 1.0 Hz which indicates the predominance of the velocity dependence of the system in that frequency range. Below 0.1 Hz, the gradient closely approximates -1 log unit / decade which indicates that for very low frequencies (i.e., for long time periods), the system then responds primarily to angular acceleration as the damping characteristics are nullified as discussed above.

For long duration or low frequency rotational stimulation, the detection threshold value is also very sensitive to experimental conditions and as indicated above is now measured in units of angular acceleration rather than in angular velocity. Clark (1970) surveyed a range of such experiments and found that the mean angular acceleration about the z -axis was 0.32 deg/sec/sec with a range of 0.05 to 2.2 deg/sec/sec.

This experimental data derived from studies on live human subjects accords well with neurophysiological investigations of mammalian end-organ response. For example, Fernandez and Goldberg (1976) investigated the response characteristics of individual neurones innervated by the sensory epithelium from the semicircular canals in Squirrel Monkeys. Such studies revealed that there were essentially two types of units called the regular and irregular units. The regular (or tonic) units had a steady resting discharge rate while the irregular (or phasic) units had an unsteady resting firing rate as their name suggests. Bode plots of the gain and phase lead of the discharge output of the two types of units as a function of the input stimulus are shown below in Figure 57.

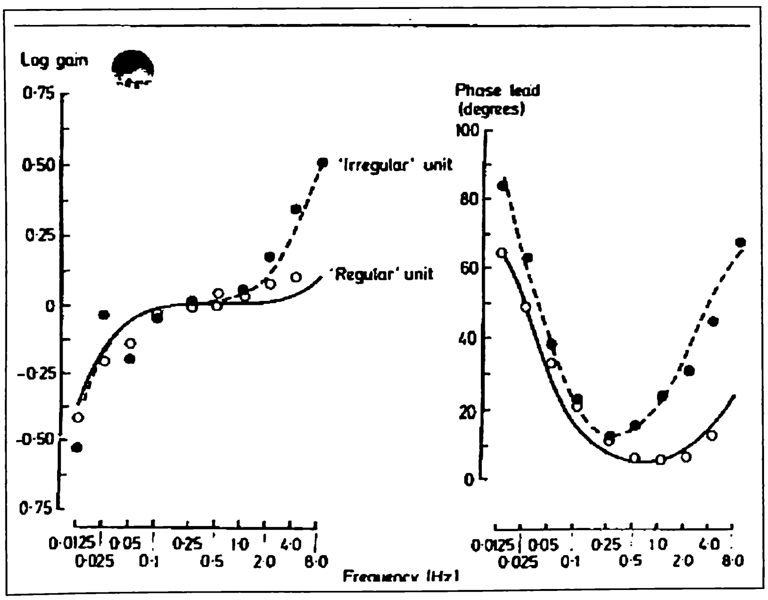


Figure 57
Bode Plots of Afferent Activity from Semicircular Canal Units in Squirrel Monkey (Data from Fernandez and Goldberg, 1971)

In Figure 57 we see that over the range 0.1 to 1.0 Hz that the gain is essentially constant and there is little phase error. This corresponds to the system signalling angular velocity. Below 0.1 Hz we see that the gain decreases indicating a reduced sensitivity and that the phase lead increases which clearly correlates to angular acceleration. Notice that above 0.5 Hz, the irregular units increase in gain and phase lead and hence become more responsive to angular acceleration whereas the regular units remain sensitive to angular velocity. This distinction allows the system to become sensitive to very short duration angular jerk while still retaining overall response to angular velocity. The upper limit of the system response is probably of the order of 30 Hz in order to account for the vestibulo-occular response performance (personal communication with Stott, J.R.R.).

Another finding of importance is that there is a fixed relationship between the level of angular acceleration experienced by a subject during sustained acceleration and their response time to detect it. Figure 58 below shows the results of 8 different studies which all have remarkably similar gradients when plotted on a log-log scale.

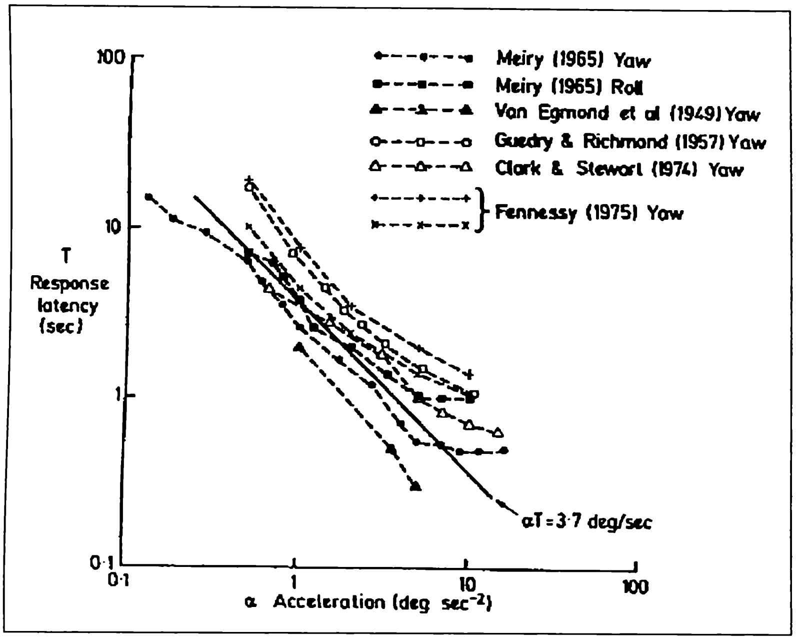


Fig 58

Relationship Between Reaction Time taken by Subjects to Detect a Step Input and the Magnitude of the Input

(Taken from Benson, 1990)

The constant gradient on a log-log plot means that the product of acceleration x time will be a constant and in this case is equal to **3.7 deg / sec**, a measure of angular velocity. This value accords well with the values shown in Figure 56 and is slightly higher because the reaction time includes both the perceptual integration of stimuli and also the mechanical integration of the canal-cupula-endolymph system.

Most natural head movements which involve changes in angular velocity are well above these detection thresholds. The passive motion of the head that accompanies walking or running commonly has peak angular velocity values in excess of +/- 10 degrees / sec at 1-2Hz. Voluntary head movements such as when a subject turns their head to fixate on an object detected in the peripheral visual field usually have a peak velocity in excess of 100 degrees / sec and may be as high as 400 degrees /sec (Benson, 1990).

6.1.3 Linear Motion Detection

The detection of linear motion is performed by the maculae which are contained in the saccule and utricle. Figure 59 below shows a schematic outline of a macula.

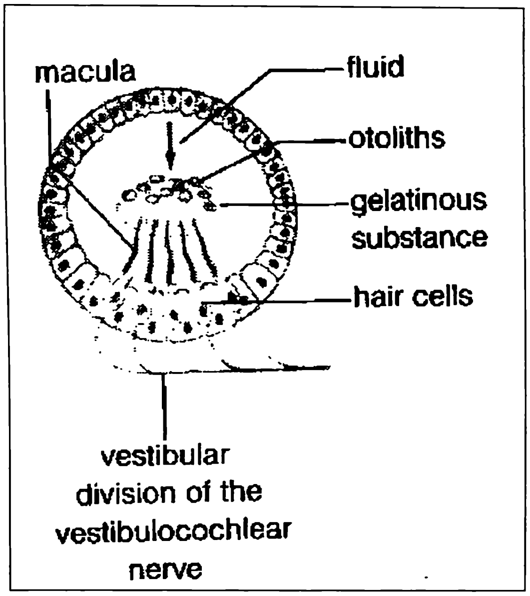


Fig 59 Macula of the Saccule (Creager, 1983)

The macula is a receptor cell which contains a gelatinous substance mounted on top of the stereocilia. Imbedded in the gelatinous cap are tiny fragments of bone called the otoliths which serve to increase the density of the cap. Because this cap has a density of approximately twice that of the surrounding endolymph, it is deflected when subjected to a linear acceleration (Graham, 1990). Unlike the semicircular canals however, the system is not heavily damped by viscous forces so the afferent discharge of the otolithic receptors approximates the linear acceleration of the head.

There is again, considerable scatter in the estimations of detection thresholds for linear motion depending on the type of experiment used. Gundry (1978) reviewed 18 sets of experiments for linear oscillation. For linear oscillation at 0.3 Hz in the horizontal plane with the subject blindfolded, the absolute detection threshold is about 0.03m/s/s and in the z plane it is double at about 0.06 m/s/s.

6.1.4 Neural Pathways from the Vestibular System

While our principle reason for discussing the vestibular system is to investigate its effect on whole body motion perception and a subject's vehicle control, it should not be forgotten that the primary role of the vestibular system is to provide information on the movement and orientation of the head which is used to regulate motor activity at the sub-cortical level. The major afferent projections from the vestibular receptors are to the cerebellum and brain stem nuclei. From these integrative centres, the information moves to the oculomotor nuclei for the control of eye movements (vestibulo-occular reflex) and to the spinal motor neurones for the control of posture and muscle activity. Some second order vestibular neurones relay to the thalamus and project to the cerebral cortex where evoked vestibular responses are localised to a small area of the post central gyrus at the lower level of the intraparietal sulcus (Benson, 1990). This lack of cortical representation of vestibular information accords with common experience that we are not normally consciously aware of the vestibular system. Most of the vestibular processing occurs subconsciously and this fact may well turn out to have important ramifications for training VDEs in the perception of vehicle motion.

6.2 Method

6.2.1 Equipment

Two types of vestibular apparatus were used;

i) The Linear motion platform

ii) The Rotary motion platform

Both sets of apparatus were housed at the RAF's Defence Research Agency Aeromedical facilities at Farnborough, England. This equipment had previously been used by the RAF for a number of published investigations (Benson and Brown, 1989; Stott, 1984; Stott and Benson, 1990; Golding and Benson, 1993; Benson and Brown,1989). The same equipment was also used for the European vestibular experiments on the Spacelab-1 mission to measure the effects of micro-gravity (Benson, Kass and Vogel 1986). The performance of this equipment has been historically well established.

a) Linear Motion Platform - Mechanical Specifications

Horizontal motion was produced by a sled mounted on a 7 metre beam fabricated from cast aluminium plate. The sled was supported on the beam by pressurised aerostatic bearings to produce a virtually noiseless and frictionless movement. The sled was enclosed by a canvas shroud to reduce any relative air motion past the subject and to reduce any visual cues. A specially fabricated aluminium bucket seat was mounted inside the canvas shroud for the subject to sit in. The bucket seat was designed and padded to distribute the accelerative forces across the subject's back, arms, shoulders, legs, and posterior to reduce local somatosensory perception of acceleration. The subject's head rested against angled and padded head supports to keep their head in a

fixed position relative to their body. The seat could be mounted either in the lateral or longitudinal position, but for all our experiments was fixed in the lateral position.
In addition to the canvas surround which was designed to reduce extraneous visual and auditory cues, the subjects wore a blindfold and insulated headphones through which broad band noise of 60 dBA was played. The subject was firmly constrained in the bucket seat by a 5 point harness with a quick release mechanism.
The sled was accelerated by an electronic servo motor controlled by an HP -85 microcomputer which provided 11 bit velocity resolution. The motion of the Linear sled was monitored by 3 independent means to ensure that the actual motion of the sled corresponded with the demand function generated by the microcomputer.
Acceleration was measured by a servo-accelerometer (Q-Flex QA -118-15) fixed to the sled carriage, velocity by a tachogenerator (Inland Type TG -2916) and position from a potentiometer directly coupled to the shaft of the servo motor. In typical operation, the system produced acceleration noise which was less than 0.01m/s/s RMS.

b) Rotary Motion Platform - Mechanical Specifications

Rotary motion was generated about the earth's vertical axis by a turntable that was driven by a precision torque motor (Artus Type MCS 1701 generating peak torque of 4.6 Nm). The motor was coupled via a rubber covered pulley to the periphery of a 0.6 m diameter alloy disc that was rigidly connected to the shaft of the turntable. The effective drive reduction was 20:1. The servo system controlling the speed of the motor was generated by an HP-85 microcomputer. The motion of the Rotary platform was measured by two independent means. The angular velocity of the turntable was monitored by a tachogenerator (Inland type 2916c) directly coupled to the shaft of the drive motor. The angular acceleration of the platform was transduced by a linear accelerometer which was mounted 2m from the axis of rotation on a stiff radial beam that was fixed to the platen of the turntable.

The platen of the turntable carried a seat and cage assembly similar to that for the Linear Motion platform as described above.

6.2.2 Motion Stimuli

During motion discrimination experiments, it is desirable to minimise any discontinuities or sharp changes in motion that the subject could use to help identify the stimuli (Benson, Hutt and Brown, 1989; Benson, Spencer and Stott, 1986). A cosine bell stimulus was therefore used because it minimises sharp changes in the next highest derivative of motion that the subject feels throughout the stimulus presentation. For the Linear tests the cosine bell was defined in terms of lateral velocity as the linear vestibular system responds primarily to acceleration while for the Rotary tests the cosine bell was defined in terms of angular displacement because the semicircular canals respond primarily to angular velocity because of the damping in the physiology and the integration in the neurology (as previously discussed).

The formula for the Base stimuli are given below;

For Linear Platform v = A1((1-Cos(wt)).Sin(wt)) velocity of motion

For Rotary Platform theta = A2((1-Cos(wt)).Sin(wt)) angular displacement

where t = 0 to 2 pi seconds

w = 1 radian / second

A1= 1 / 1.2959 giving peak lateral acceleration of +/- 2.11 m/s/s

A2 = 60 / 1.2959 giving peak angular acceleration +/- 126.65 degrees/s/s

Thus each stimulus had a duration of 2 * pi seconds (6.28 seconds). Both these stimuli are at least 30 times larger than the minimum that most subjects can detect according to the preceding discussion.

Figures 60 to 62 below show the displacement, velocity and acceleration profiles of the Base motions for the Linear platform while Figures 63 to 65 show the corresponding Base motions for the Rotary platform.

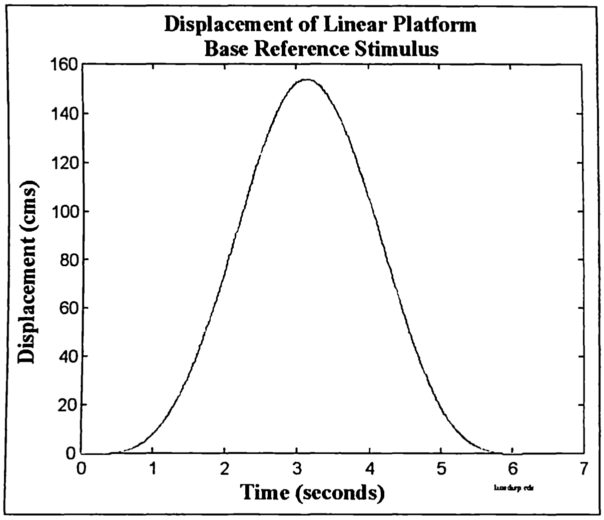


Figure 60

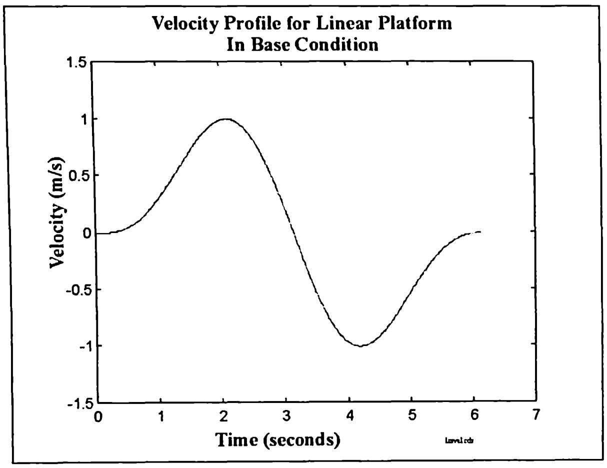


Figure 61

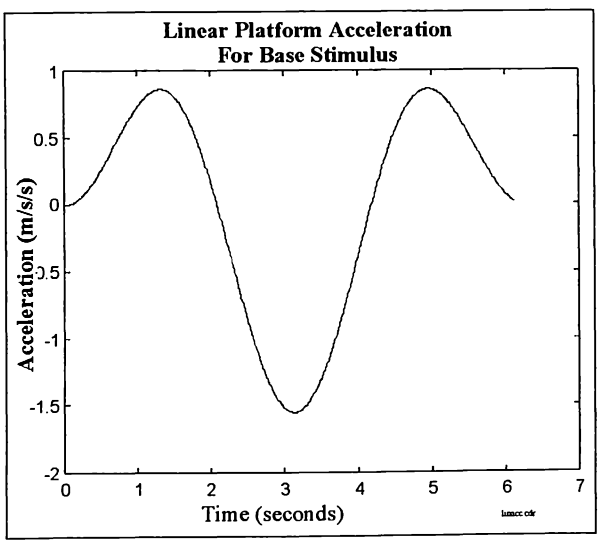


Figure 62

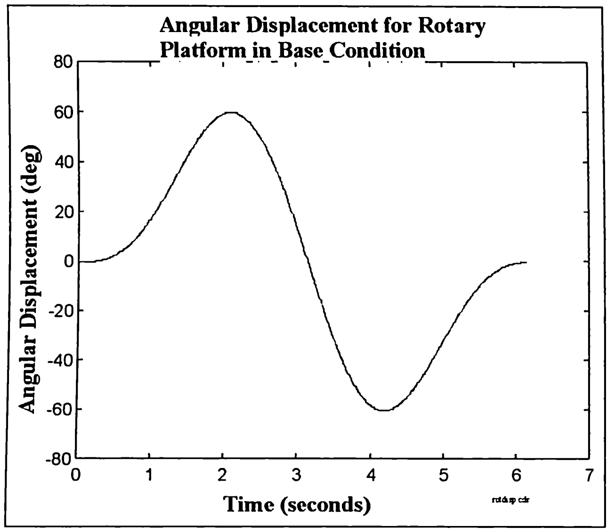


Figure 63

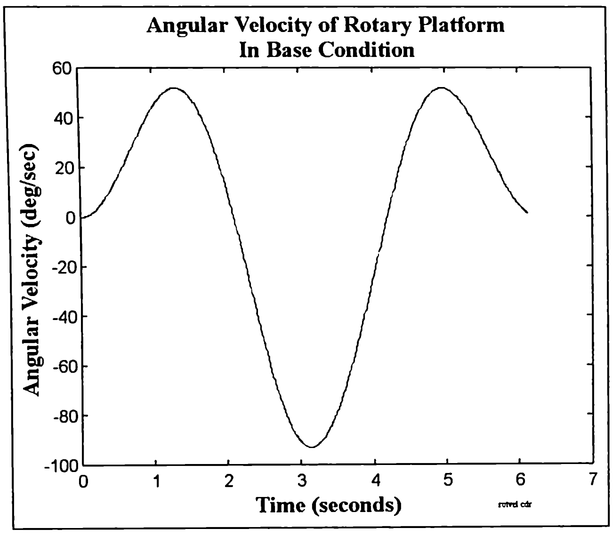


Figure 64

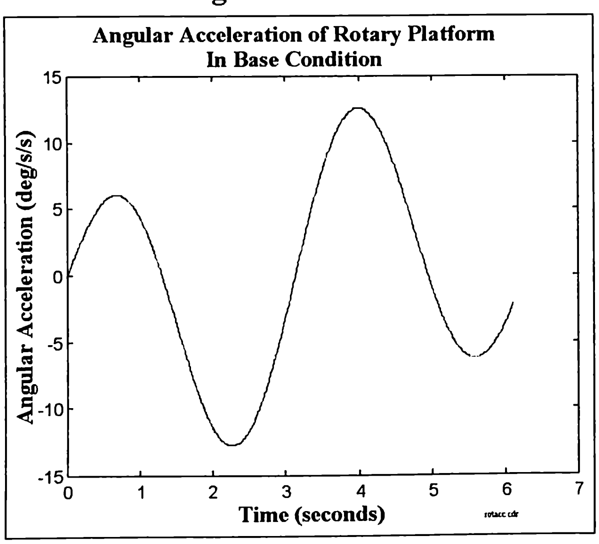


Figure 65

For the Embedded Discrimination test the Base stimulus was constructed by superposing a perturbing wave within the standard waves described above.

The perturbing waves were identical to the standard waves except that the amplitudes and frequencies were scaled. For the Rotary test the amplitude was ¼ of the Base wave amplitude. For the Linear test the amplitude was ½ of the Base wave amplitude. In both cases the frequency was twice the Base frequency and the point of insertion within the Base wave was at t=2.61 seconds. Figure 66 shows the construction of the velocity wave for the Lateral Embedded test.

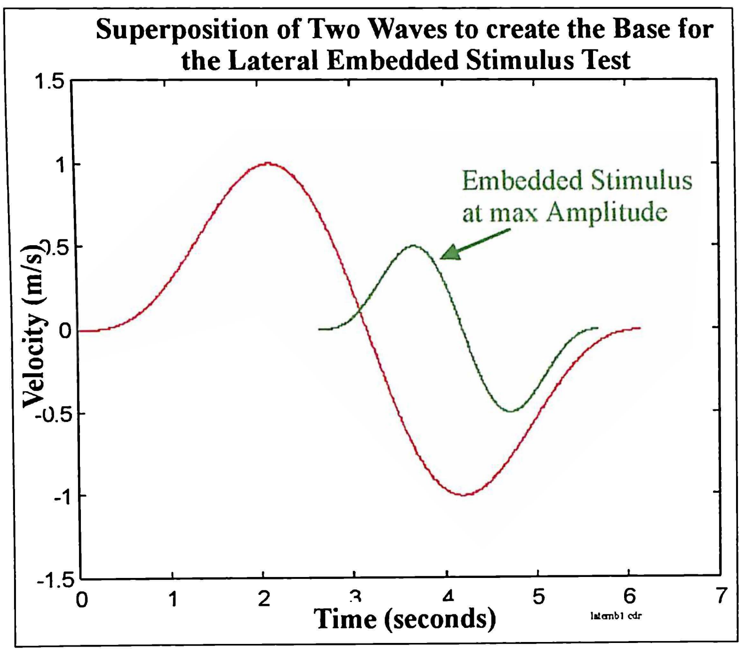


Figure 66

In addition to these standard stimuli (Figures 60-66) subjects were also presented with attenuated stimuli. For the Amplitude Discrimination test the entire waveform was attenuated while for the Embedded Discrimination test only the perturbing stimulus that was superposed was attenuated. The attenuation difference between the Base and the Modified stimuli was expressed in dB where

$$dB = 20\ Log_{10}[\ (\ A_{base}\text{-}A_{modified})/A_{base}]$$

Figure 67 shows a comparison of the velocity profiles for the Base stimulus and a -10dB Modified stimulus for the Lateral Amplitude test.

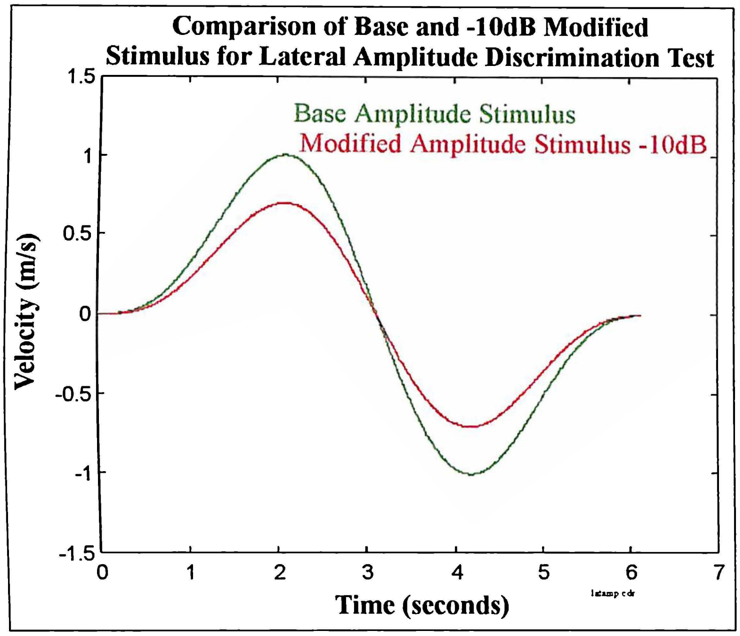


Figure 67

Figure 68 shows a comparison of the velocity profiles for the Lateral Embedded Base Stimulus with a -10dB Embedded Stimulus.

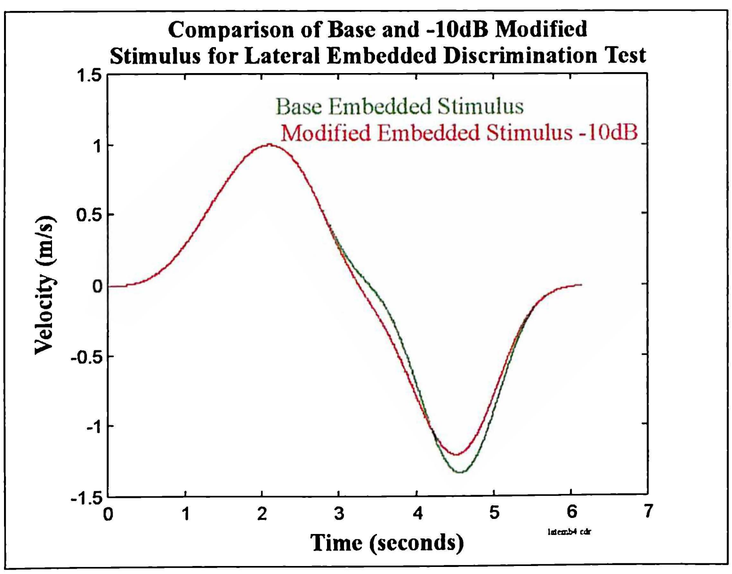


Figure 68

Figure 69 shows the Lateral acceleration corresponding to Figure 68.

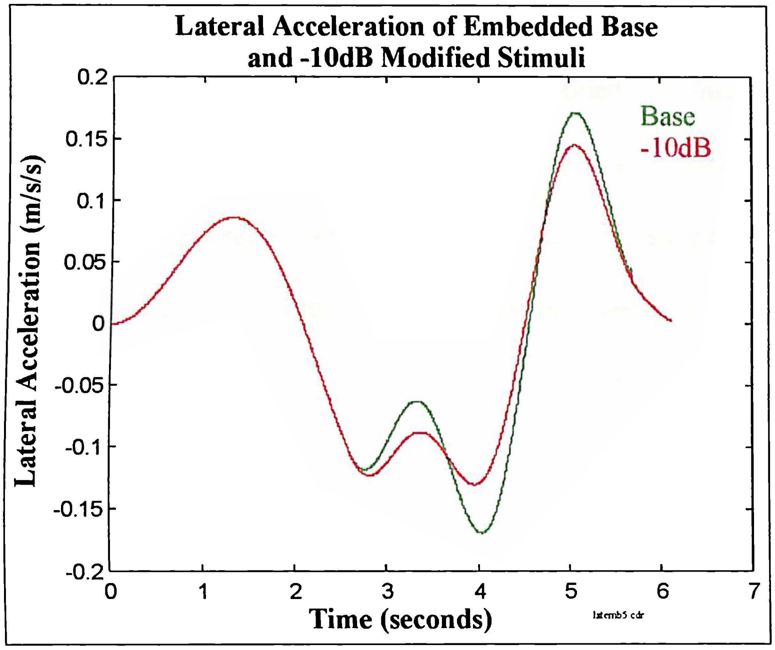


Figure 69

6.2.3 Subjects

Experiment 1: Six non pilot RAF staff aged 20 -32, 4 Males and 2 Females.

Experiment 2: Twenty four Ford VDEs aged 24 - 56, 22 Males and 2 Females.

6.2.4 Procedure

Two sets of experiments were conducted;

a) Experiment 1

Experiment 1 was conducted as a pilot study to investigate the correlation between the thresholds established under the 4 experimental conditions;

i) Rotary Embedded Discrimination

ii) Rotary Amplitude Discrimination

iii) Linear Embedded Discrimination

iv) Linear Amplitude Discrimination

and further to investigate the repeatability of these measurements over 5 successive days. Each subject's discrimination thresholds were measured on all 4 discrimination tests (Lateral, Rotary, Embedded and Amplitude) 5 separate times over 5 successive days. Thus each subject was tested 20 times over the same 5 days. Subjects were tested as a group, one after the other, until all subjects had completed their first test. Subjects were then successively tested on their second test. This process was repeated until all subjects were tested on all 4 tests in the same day. Subjects rested while the other subjects were being tested. The order of presentation of the 4 tests was randomised between the subjects but kept constant for each subject on successive days. Each subject was re-tested on the same test at the same time of each day (+/- 5 minutes) to reduce the influence of diurnal factors.

b) Experiment 2

Experiment 2 was conducted to investigate the effect of different instructions on subjects' Rotary Amplitude Discrimination thresholds and the correlation of these thresholds with the subjects' Lane Change performance.

Subjects were tested individually on separate days. Each subject was tested 4 times on only one test, the Rotary Amplitude test. In the first test subjects were given no instructions on how to perceive or encode the platform motion. This was termed the 'Free' condition. In the following 3 tests, the subjects were successively instructed to perceive the motion by visualising the displacement, velocity and acceleration of the platform

For both sets of experiments, subjects were presented with 30 pairs of trials, each consisting of a Base and an attenuated Modified stimulus. Subjects were required using the 2 alternate forced choice procedure to determine which of the pair was the attenuated stimulus. The subjects responded by use of a response box which was held on their knee with both hands. The response box contained 3 buttons by which they signalled their answers;

i) Left Button to select first stimulus presentation

ii) Right button to select 2nd stimulus presentation

iii) Large Central Button Emergency Stop.

A microphone was mounted on the headphone assembly so that the subjects could communicate to the experimenter at any time. In return, the experimenter could override the broad band masking noise in the subject's headphone and communicate with the subject. All subjects were given 5 practice trials prior to testing.

The attenuated stimulus level was set at -20dB for the first trial and then adjusted using the QUEST algorithm on subsequent trials (Watson and Pelli, 1983) to determine the subjects' discrimination thresholds. A Monte Carlo Simulation had been previously conducted to find the best number of trials to use with the QUEST algorithm given the psychometric curves we had obtained with during preliminary tests. The Monte Carlo simulations indicated that 30 trials would be suitable for determining the subjects' thresholds to the accuracy we required. This topic is addressed more fully in section 6.7 as it turns out to play a pivotal role in the conclusions that we can draw from these experiments.

6.3 Experiment 1 Results

The correlation coefficients (Spiegel, 1961) relating the scores between the 4 different tests collapsed across days is given below in Table 3.

	Linear Amplitude	Linear Embedded	Rotary Amplitude	Rotary Embedded
Linear Amplitude	1			
Linear Embedded	-0.002	1		
Rotary Amplitude	0.02	0.02	1	
Rotary Embedded	0.03	0.2	0.05	1

Table 3
Vestibular Threshold Correlations
For Internal Subjects

A sampling theory of correlation analysis for the r values contained in Table 3 revealed that there was no significant relation between any of the tests. These null results could mean;

A) The tests measure orthogonal attributes and therefore Linear, Rotary, Amplitude and Embedded Thresholds are not related.

or

B) The tests have poor inter-trial reliability. That is, a subject's threshold on say the Linear Amplitude test on one day is not related to their threshold on another day.

To investigate hypothesis B, we plotted each subject's threshold estimates for each individual test over the 5 successive days that they were tested. Figure 70 below shows the results obtained in this manner for the Linear Amplitude test.

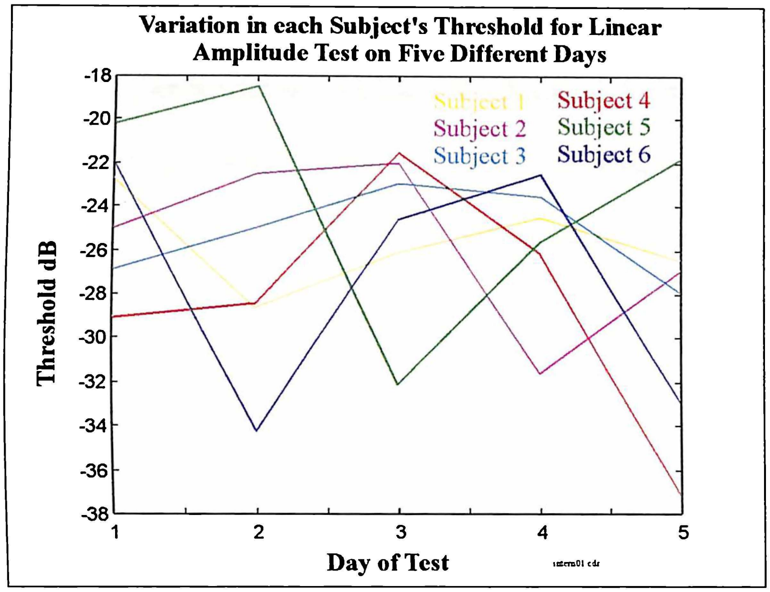


Figure 70

Similar results were obtained for the Linear Embedded, Rotary Amplitude and Rotary Embedded tests. It is immediately obvious from Figure 70 that inter-test reliability is of considerable concern. For example we see on days one and two that subject 5 performs the worst but on day 3 performs the best. Similarly, subject 6 obtains the 2nd to worst result on day one but the best result on day two only to revert to being the worst performer on day four. It thus appears that hypothesis B is true, namely that the thresholds vary from day to day.

This raises the question as to why the threshold estimates were so variable. There are 3 hypothesis that could account for this;

i) The equipment did not perform reliably from one day to the next which meant our estimate of each subject's threshold was inaccurate because the stimuli we presented did not equate to the stimuli we thought we were presenting

ii) We did not get an accurate measure of each subject's threshold on each day because either we did not present enough stimuli to the subject or we did not present the right level of stimuli to the subject. That is, the subjects' thresholds were well defined but we did not sample correctly to estimate them. This would mean that the error bars for each measurement in Figure 70 were large compared to the differences in threshold measured on successive days or between subjects

iii) The subjects' thresholds actually do change on a day to day basis

Inspection of the telemetry data which recorded the actual motion of the platforms during each test indicated that the equipment performed correctly. The target motions and the actual motions varied by less than 0.5dB for all trials and all subjects. Thus hypothesis i) can be eliminated.

In order to estimate whether the subjects' thresholds were reliably determined, 2 sets of preliminary analysis were undertaken (this topic is treated more formally in section 6.7 as this question reappears and is central to the interpretation of the main vestibular study, Experiment 2).

In the first analysis we determined the mean inter-trial variance of the Weibull estimate (Weibull, 1951) of each subject's threshold as a function of trial number. After each trial, the QUEST algorithm produces a likelihood function which estimates the probability density of the threshold for various stimulus levels. The mean variance of the Weibull estimate on successive trials, across all subjects, is a reasonable indicator of the reliability of the threshold estimate (from our Monte Carlo simulations). Figure 71 below shows this result for all subjects from Experiment 1.

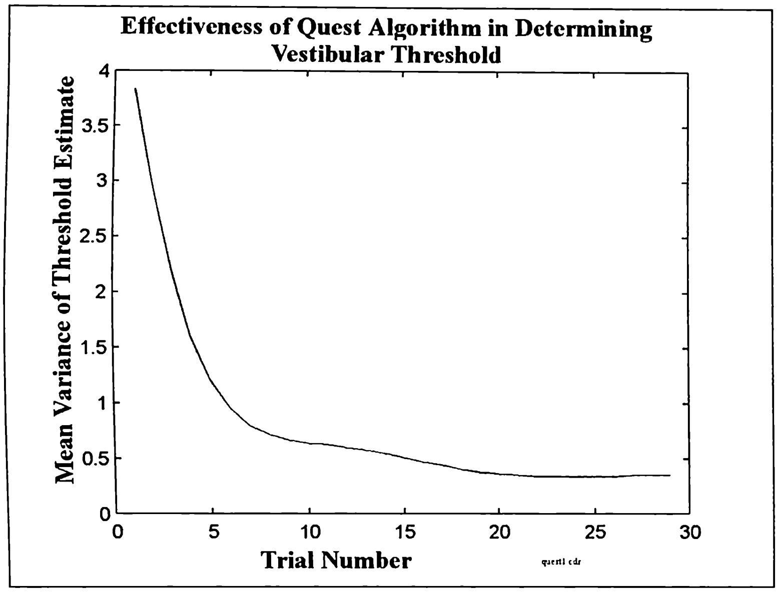


Figure 71

Figure 71 shows that after trial 20, the Quest algorithm has converged to a threshold estimate which has a mean error of approximately +/- 0.5dB on successive trials. This indicates that the estimates of each subject's threshold are likely to be determined to an accuracy of about +/- 0.5dB. If the error bars in Figure 70 are +/- 0.5dB, then the null hypothesis that there is no significant difference in threshold estimates on different days is rejected at the p=0.001 level. That is, it is very likely that the subjects actual thresholds varied from one day to the next.

The second method for estimating whether the subjects' thresholds were actually obtained, was to observe the probability of the subjects' correct responses at the calculated threshold value. The threshold is after all defined in terms of a probability point on the psychometric curve. In the QUEST algorithm, the assumed form of the psychometric function is a Weibull function. This is defined as;

$\omega t(x) = 1 - (1 - \gamma)\ \exp[-10(\beta/20)(x - t - e)]$ (Watson and Pelli, 1983.)

Altering the value for e determines where on the psychometric function the threshold is defined. With e= 0, the threshold is defined as that stimulus level where the probability of being correct is 0.816. This canonical value also corresponds to the part of the Wiebull function which has the greatest gradient. For our trials we used a value of e= -2.3 which gives a probability of being correct of 0.663. We chose this lower value because we wanted to present stimuli at the more difficult end of each subject's psychometric discrimination curve. Our experience with VDEs in practical tests indicated that they often attempted to make judgements that were in this region of lower probability.

To obtain an estimate of each subject's probability of being correct at the claimed threshold level, we observed the number of correct responses that each subject made over their last 10 trials of each test. To increase the statistical base for this, the following graphs were produced using both the current data and the data from Experiment 2. Obviously, the actual stimulus values presented to the subjects during the last 10 trials (i.e. from trials 20 to 30) were not all at exactly the Quest's final threshold estimate, but from Figure 71 above, we see that the variance from the final threshold estimate was small. If each subject's threshold was accurately determined by the QUEST algorithm we would expect on average that the probability of each subject being correct in their last 10 trials would closely approximate our theoretical value of 0.663 at which we defined our threshold. The actual observed value was 0.62 over all the subjects providing close agreement. Finally, a Monte Carlo simulation was conducted to determine the expected frequency distribution of correct responses over the last 10 trials for all the subjects on the basis that their thresholds were indeed determined to an accuracy of +/- 0.5dB. Figure 72 shows the results of this expected distribution together with the actual distribution obtained.

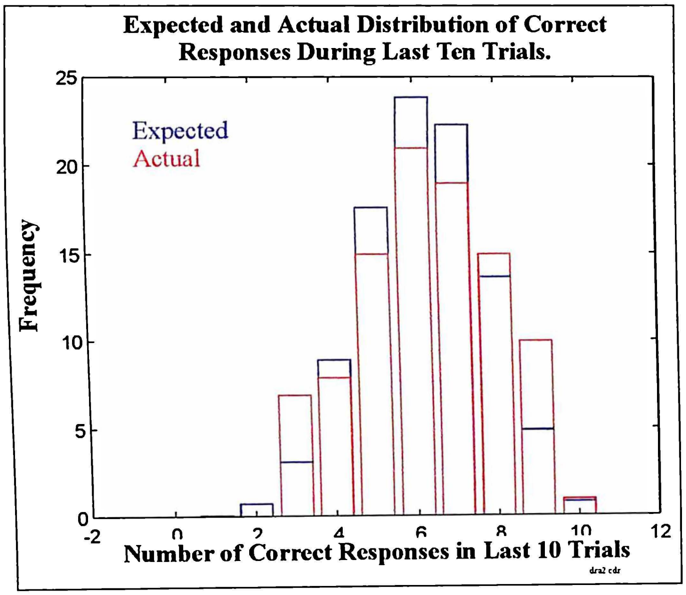


Figure 72

6.4 Experiment 1 Discussion

It would appear that the subjects' thresholds in the Rotary, Lateral, Embedded and Amplitude tests were reliably obtained on 5 successive days, but that these thresholds were significantly different on different days. Informal comments from the 6 subjects indicated that they often changed their strategy for discrimination between tests. For example, in one test a subject might concentrate on their perception of the rotary acceleration of the platform, while in another test they might base their decision on their perception of maximum angular velocity. The subjects reported that different strategies appeared to influence their discrimination performance. That is, their discrimination was not solely dependent on their physiological hardware but for example on what they attended to and how they interpreted the motion. This was quite an exciting possibility as it might mean that subjects could be taught to improve their discrimination performance by some form of training. If we could find the optimum strategies for performing the discrimination tests and train subjects to use those strategies efficiently, there existed the real possibility that improvements in their performance during the vestibular apparatus tests might carry over into their work with vehicles. Unfortunately there was little time between testing the 6 RAF subjects and the arrival of the 24 Ford subjects so we could not conduct further internal experiments on different strategies and their effectiveness.

The results from the RAF pilot study suggested that it might be profitable to modify our planned method of testing the Ford subjects. Rather than measuring each subject's threshold as if it was a fixed entity and then correlating those thresholds with some measure of their driving performance, we decided to use the vestibular test equipment to see if we could find methods for improving their discrimination performance. From the tests conducted on the RAF subjects, it appeared that the subjects' methods of visualising motion might have an important impact on their performance. At this stage we had no information on the best procedure for manipulating the subjects' visualisation and so we simply decided to test each Ford subject 4 times and instruct them to perform each discrimination test using only;

i) Displacement perception

ii) Velocity Perception

iii) Acceleration Perception

iv) Any method that they chose (Free test)

Because we were going to test each Ford subject 4 times under 4 different instruction procedures, we did not have time to test them on both the Linear and Rotary Platforms nor under both the Amplitude and Embedded stimulus procedures. That would have required testing each subject 16 times which would have taken about 6 hours of experimentation per subject and introduced fatigue considerations. Instead, we decided to restrict the four sets of instructions to just the Rotary Platform under the Amplitude Discrimination protocol. In addition to recording the subjects' results, we also decided to keep a log of their comments during each test. We hoped that the subjects' introspective comments would give us a further insight into the factors that affected their vestibular performance.

6.5 Experiment 2 Results

The results of the Rotary Amplitude Discrimination test under the 4 different sets of instruction are shown below in Table 4.

Subject	Free	Displacement	Velocity	Accel
1	-28.72**	-25.13***	-30.08***	-20.08
2	-19.86	-27.92	-23..68	-21.06
3	-18.94	-18.15	-22.21	-30.08*
4	-17.30	-30.79***	-21.51	-14.78
5	-21.40	-26.99***	-29.73*	-27.92*
6	-24.33	-20.51	-27.45*	-21.86
7	-23.94	-28.56***	-22.33	-21.86
8	-22.57	-19.55	-35.03*	-22.81
9	-20.51	-30.0***	-22.81	-27.00
10	-21.06	-18.25	-20.73	-33.31*
11	-23.94	-22.93	-18.84	-23.05
12	-21.97	-21.97	-27.61*	-20.84
13	-21.74	-28.07***	-24.46	-23.18
14	-20.08	-17.21	-25.41	-23.18
15	-24.73	-28.23***	-21.17	-22.33
16	-23.30	-25.83***	-34.36**	-30.61**
17	-31.91**	-24.32***	-19.87	-20.95
18	-21.86	-28.72***	-16.39	-21.97
19	-29.73*	-25.13	-28.56*	-34.36*
20	-19.86	-21.74	-18.25	-22.33
21	-29.90*	-28.89	-22.45	-32.70*
22	-19.45	-28.72***	-21.06	-21.74
23	-25.41	-22.93	-31.54	-20.73
24	-21.97	-22.09	-19.14	-23.30
Mean	-23.10	-24.70	-24.36	-24.25
Std Dev	3.8	4.1	5.1	4.8

Table 4
Vestibular Threshold Results for Ford VDEs

* = results where subjects reported concentrating their attention on the maximum amplitude (Focus Restriction)

** = results where subjects reported using a Special Encoding scheme

*** = results where subjects reported using an Arc Scribing technique

The Student's t-test (Spiegel, 1961) revealed that there was no significant difference in discrimination thresholds between the 4 sets of instructions (Free, Displacement, Velocity and Acceleration) at the $p = 0.1$ level of significance.

a) There was a highly significant difference ($p < 0.001$) in discrimination performance between the trials where subjects reported restricting their focus of attention to the maximum amplitude of the stimulus (Focus Restriction, mean = -30.2, std 2.5) and the trials where subjects reported no specific technique of discrimination (mean = -21.8, std 2.8).

b) There was a highly significant difference ($p < 0.001$) in discrimination performance between trials where subjects reported using a Special Encoding scheme (mean = -30.9, std = 2.4) and the trials where subjects reported no specific technique of discrimination (mean = -21.8, std 2.8).

c) There was a highly significant difference ($p < 0.001$) in discrimination performance between trials where subjects reported using an Arc Scribing technique for displacement (mean =-30.9, std =2.4) and trials where the subjects simply attempted to judge the angular displacement (mean =-23.8 , std =4.3).

d) There was no significant correlation ($p > 0.25$) between any of the vestibular tests and the subject's performance in the Lane Change Discrimination tests (section 8), either before or after training.

6.6 Experiment 2 Discussion

The independent variable that we attempted to manipulate in experiment 2 was the subjects' methods of visualising the platform motion. It quickly became clear that our instructions to concentrate on displacement, velocity, or acceleration were unsuccessful in achieving visualisation control. Within each of these 4 instructions, there was considerable variation as to how the subjects actually visualised the platform motion. For example, subject 16 visualised the motion of the platform in terms of his golf swing for the Velocity and Acceleration protocols (Special Encoding scheme). As the platform swung left, stopped, reversed direction and swung right, he imagined he was teeing off and these motions corresponded to his back-swing and drive. In the Velocity test he imagined the speed of his hands while in the Acceleration test he imagined the force he was using on the handle. Subject 16 reported that it felt more natural for him to modulate the speed of his golf swing than the sensation of force in the handle. For the Displacement test subject 16 imagined the platform scribing an arc along the ground and for the Free test he used his 'general impression of motion'. Subject 16 was a competitive golfer and had played golf every week for the past 12 years.

Subject 1 on the other hand, used music to encode the Free and the Velocity tests, the Arc Scribing scheme for the Displacement test and no particular encoding method for Acceleration test. In the Free and Velocity tests, subject 1 converted his perception of the angular velocity of the platform into a musical pitch. The faster the platform appeared to rotate, the higher the pitch. After each such trial, subject 1 had a musical tune in his head which corresponded to the motion of the platform. He performed the

Discrimination test in these cases by comparing the musical tunes. Subject 1 had played in a German orchestra for the past 17 years.

Thus, although subjects 1 and 16 received the same instructions, their method of visualising the motion was quite different. Furthermore, subject 1 used his Special Encoding scheme on the Free and the Velocity trials while subject 12 used his Special Encoding scheme on the Acceleration and Velocity trials. That is, not only were the encoding schemes different, they were also applied to different test protocols.

In the trials where subjects reported restricting their attention to the portion of the waveform where the maximum amplitude occurred, the subjects simply concentrated on how large the biggest part of the waveform (either velocity or acceleration) felt to them. After the largest part of the waveform had passed, they tried to retain an accurate impression of that motion and ignore the subsequent motion. Their discrimination was based on a comparison of their impression of the two maximum amplitudes.

Thus our formal instructions did not appear to adequately control the subjects' methods of motion visualisation in a systematic fashion. Despite our instructions, the subjects were still able to employ a variety of visualisation techniques and these uncontrolled variations appeared to have had a stronger influence on their discrimination performance than our instructions did. If so, this would account for the failure of the 4 instruction sets to produce a significant difference in discrimination performance. Nevertheless, because we logged the subjects' introspective comments, we were able to a- posteriori divide the trials into 4 classes;

i) No particular method was reported

ii) A Special Encoding scheme (music or golf) was reported

iii) Subjects reported restricting their focus to only the section of the waveform containing the maximum amplitude (Focus Restriction)

iv) Subjects reported using the Arc Scribing technique

As the results section 6.5 above showed, visualisation techniques b), c) and d) produced highly significant differences in performance over the unspecified visualisation techniques.

The musical technique arose during the Lane Change program after subject 1 had completed his first Lane Change Discrimination tests where he had scored at chance level on the easiest setting (20 psi). During the Training session that followed the Discrimination test, subject 1 reported having difficulty remembering the motion of the vehicle after each run. He felt that his memory for the vehicle's motion gradually faded after each run so that by the time the 2nd run of each pair was completed, he had forgotten what the first run felt like and so could not discriminate between them. Inspection of the telemetry data revealed that subject 1 was controlling the vehicle well and that there were large systematic differences in the objective behaviour of the vehicle between the Base and the Modified conditions. That is, the telemetry indicated he should have been able to easily discriminate between the vehicle settings. Considerable discussion took place with subject 1 about techniques for improving his memory. One of the items covered was the importance of mental structures for interpreting and encoding incoming events. We discussed the performance of Grandmaster chess champions who could encode and remember entire chess board layouts in matters of seconds with apparent ease. It was found that they could do this only if the chess pieces were in legal positions. If the chess pieces were placed in positions that were not possible, then the Grandmasters' performances fell to approximately the same level as normal chess players. The implication is that the Grandmasters have large and well developed encoding schemes. Rather than remembering the individual pieces on the board, they appear to remember patterns of threat and defence where entire sections of the chess board can be summarised as a single complex interplay of pieces. Thus the incoming information is interpreted using a rich framework that has a well developed syntax and grammar.

Subject 1 was an accomplished musician and so he wondered whether he could use his well developed musical framework to help encode the motions he was experiencing. Unfortunately, because we had already spent considerable time trying other training procedures with him (to no apparent effect) there was insufficient time to test this hypothesis in this Lane Change session. The following day, subject 1 visited Farnborough for his vestibular tests. He tried his musical encoding scheme during the Free and Velocity tests (-28.75dB and -30.08dB) by encoding angular velocity as musical pitch. During the Acceleration test he reverted back to the techniques he used during the Lane Change test and his performance fell to -20.0dB. Two days later, subject 1 was again tested using the Lane Change Discrimination test without any further training. He decided to use the musical encoding scheme for this test. There was a remarkable change in performance. Whereas previously subject 1 could not discriminate between the easiest settings (20 psi vs 36 psi), with the new encoding scheme he scored 100% correct on every single trial right up to the maximum setting (34 psi vs 36 psi). This was a level of performance previously only demonstrated by

Jackie Stewart. Thus, not only did the visualisation technique appear to produce an improvement in the vestibular tests, it also appeared to carry over into practical vehicle evaluation.

A similar situation occurred with subject 16, the only difference being that subject 16 was an accomplished golfer and visualised the motion of the vehicle or platform in terms of his golf swing.

The idea of concentrating attention on only one part of the stimulus motion also arose during Lane Change Discrimination training. As we shall see later, altering the tyre pressure settings produced the greatest difference in vehicle behaviour in the transition section of the Lane Change with smaller changes occurring elsewhere. It was found by trial and error that getting the subjects to restrict their attention to just the transition section produced an improvement in discrimination performance. The interesting point about this attention restriction was that the improvement in discrimination performance occurred even when the subjects already knew that the main difference was occurring in the transition section. It wasn't so much that they were being directed to attend to a stimulus event that they had been previously ignoring, rather they were being instructed to ignore additional stimulus events occurring outside the main difference. That is, the subjects were already deriving most of their information from the transition section and simply obtaining supplementary information from the rest of the Lane Change in an attempt to assist them in their discrimination. It appeared that deriving this supplementary information was somehow interfering with their perceptual ability at the most important part of the Lane Change. (We shall return to this topic in greater detail in section 7).

Because this process appeared to the subjects to be successful in the Lane Change test, it was also tried by some of them in the vestibular tests with the significant results detailed above in section 6.5.

Subjects reported that the Arc Displacement visualisation technique appeared to function in a manner that was similar to a combination of the Focus Restriction technique and the Special Encoding technique. What most subjects attempted to do was imagine a room with objects scattered around or a circular wall with markings on it. As they were rotated they imagined moving past these objects until the chair stopped and reversed direction. At that point, they simply noted what object was in front of them or where they were relative to the imaginary markings on the wall. Thus the imagery they developed was to some extent similar to having a Special Encoding scheme. They generated their own set of imaginary objects fixed in space which they used to help encode the motion stimuli they were experiencing . They could even re-calibrate these imaginary objects by turning their head through 90 degrees while the chair was stationary between trials. Secondly, once they had encoded the position of furthest rotation, they would ignore the rest of the motion. To this extent it resembled the Focus Restriction technique.

It was interesting that the three techniques, Special Encoding, Focus Restriction and Arc Scribing all produced statistically the same results (means = -30.2, -30.9, -30.9 and std = 2.5, 2.8, 2.4). It may be that employing these techniques had activated some common psychological processes. For example it may have simply been that using these techniques forced the subjects to attend to the stimuli more or that they were required to process them to a greater level.

The failure to establish any correlations between the vestibular tests and the Lane Change tests could have arisen because there actually is no relation between the two skills or more likely it could have arisen because of methodological considerations. The order in which the subjects were presented with the Lane Change Discrimination tests and the vestibular tests was randomised between subjects. Some subjects received their Lane Change instruction prior to the vestibular tests while others received it afterwards. It was only during the Lane Change tests that the subjects received tuition on methods of improving perception. Thus some subjects had received perceptual training prior to the vestibular tests while other subjects hadn't. If vestibular and Lane Change performance **can** be affected by training, then the lack of correlations could well be explained by the random application of training across these two tests.

Unfortunately, we were not able to investigate these special perception techniques (Focus Restriction, Arc Scribing and Special Encoding schemes) in more detail or in a controlled fashion during the vestibular tests owing to the organisation of the program in England and the overriding constraints imposed on us by Ford Motor Co. The primary purpose of the VDEs visit to England was to improve their vehicle evaluation performance as much as possible. VDE time was jealously guarded by Ford management and this coupled with the high cost of undertaking the VDE training and evaluation program (in excess of US$ 1 million) meant that sometimes experimental formality was subsidiary to commercial constraints. Ford management made it quite clear that the only alternative to these commercial pressures was to attempt the research without the use of professional VDEs and without Ford's support - both financial and facilities. It was for these reasons, for example, that no control subjects were used in the Lane Change training modules. This situation is similar to that faced by drug companies trialling a new drug for people with very serious illnesses. Previously, researchers were required to include the use of placebos in double blind trials to prove the effectiveness of a new drug but in recent years ethical considerations have led to a marked change in research procedure. In particular, the Declaration of Helsinki ratified in 1989 (Weijer and Elliott, 1996) says that all patients enrolled in a research protocol, including those in a control arm, must be assured of the best available treatment. Just as researchers can not let subjects die untreated to establish the control level, we also could not fail to give each subject the best training that was available to us. Nevertheless, our record of each subject's introspective comments during their vestibular tests allowed us to separate out the 4 classes of visualisation techniques used separately from that imposed by our instructions. This class separation showed that the same techniques that improved performance in the Lane Change also appeared to produce significant improvements in the vestibular test.

Even although one is reluctant to draw too strong a conclusion from these limited studies of 24 Ford subjects and the 6 DRA subjects, the results seem to suggest that supra-liminal vestibular discrimination performance may not primarily be determined by innate physiological processes for most VDEs. While there may well be some upper limit of performance imposed by each subject's physiology, it would appear that for most subjects, higher cognitive processes are currently the limiting factors and that there is significant scope for improvement by manipulating them. This is a most exciting proposition for the vehicle dynamics community and as result Ford Motor Co has commissioned the development of what has been loosely termed a 'Motion Gymnasium' for its VDEs.

6.7 Statistical Analysis of Vestibular Results

6.7.1 Importance of Threshold Error Estimates and Development of a new Statistical Algorithm

In section 6.3 we briefly looked at the issue of whether or not we had reliably determined each subject's vestibular thresholds in Experiment 1. That preliminary analysis indicated that we had done so to an accuracy of approximately +/- 0.5dB. However, because the strength of our conclusions in Experiment 2 are so dependent on our confidence that the thresholds were reliably determined and because these conclusions have such powerful ramifications, it was most important to ascertain with greater mathematical confidence the size of the threshold uncertainties.

McKee, Klein and Teller (1985) had performed extensive computer simulations to study the accuracy of probit analysis estimates of threshold estimations for psychometric functions derived from 2 alternate forced choice experiments. Their study concluded that standard probit analysis equations do not give valid estimates of threshold variability. In all cases the actual threshold errors were greater than the errors reported by probit analysis. It was with this in mind that I approached a professional statistician who specialised in psychometric curves to assist me in obtaining a better estimate of the error in the vestibular thresholds recorded in section 6.5 above. This consultation lead to the development of a new algorithm tailored specifically to the 2 alternate forced choice paradigm which gives reliable estimates of the threshold errors given the type of data we obtained. The theoretical derivation and justification for this algorithm based on an inverse link function is beyond the scope of this thesis, properly falling within the discipline of statistics and hence is not presented here. However, a brief description of the algorithm (tentatively referred to as the Jorgensen algorithm after its author) is presented below in section 6.7b.

6.7.2 Outline of the Jorgensen 2 Alternate Forced Choice Algorithm

1) Assume

k_0 is the initial gradient of the psychometric curve at the initial estimate of the threshold t_0

2) Calculate the parameters a_0 and b_0 which describe the psychometric function from k_0 and t_0 by;

$b_0=k_0 /(2p-1)(1-p)$ where p = probability of being correct at the threshold

$a_0= - bt_0$

3) Let

x_i be a vector containing the magnitudes of the i stimulus presentations

y_i be a vector containing either 1's (if the subject was correct on that trial) or 0's (if the subject was incorrect on that trial)

from this form the column vectors

$q_{0i} = a_0 + b_0.\ x_i$

$p_{0i} = (1+ e.\text{^}(q_0)i)./\ (2+e.\text{^}(q_0)_i)$

$z_i = (q_0)_i + (yi - (p_0)_i).\ /\ (2(pi_0)_i\text{-}1).(1\text{-}(p_0)_i))$

4) Define a column of weights

$w_i = (2(p_0)_i\text{-}1).\text{^}2.\ (1\text{-}(p_0)_i)./(p_0)_i$

5) Form the augmented matrix X from x_i by adding a column of 1's equal to the length of x in the first column and placing the x_i in the 2nd column

6) Form the rectangular diagonal matrix W from the column vector of weights w_i

7) Calculate new estimates a_1 and b_1 of a and b by finding the 1 X 2 vector

$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = inv(X^TX)XWz_i$$

8) From a_l and b_l calculate new estimates of the threshold and slope at the threshold by the inverse of 2) above

9) Form the 2 x 2 matrix V from the transpose of X and the inverse of the products as follows

$V = inv(X^TWX)$

10) The new estimate of the error in the new threshold et(j), is calculated from the first, second and fourth elements of the matrix V (V_{11} V_{12} and V_{22}) and the new estimates of a and b (a_j and b_j) by summing the vector product over i

$et(j)=\sum sqrt((1./b_j.\hat{}2).*V(_{11}) + (a_j.\hat{}2./b_j.\hat{}4).*V(_{22}) -2.*V(_{12}).*(a_j./b_j.\hat{}3))$

11) The new estimates of a and b are fed back into 2) and the iteration repeated until a and b no longer change

From these values of x_i, y_i and a and b it is also possible to generate the likelihood surface which shows the probability that each value of a, b, k and t are the correct solution to fit the data presented.

The likelihood L is given by

$l = Log(L) = y_{ji} . log(p_{ji}) + (1-y_{ji}).log(1-p_{ji})$ where j is the iteration number

6.7.3 Results Using Jorgensen Algorithm on Vestibular Experiment 2

The Jorgensen algorithm converged rapidly (usually in less than 8 iterations) and produced estimates of the psychometric curve, the threshold, the error in the threshold, the gradient at the threshold, the error in the gradient and the likelihood surfaces in both the parameter planes and the threshold - gradient planes. For example, let us consider subject 22 from Table 4. Subject 22 used the Arc Scribing technique (-28.72dB) for the Displacement trials but did not report any special techniques for the Free (-19.45dB), Velocity (-21.06dB), or Acceleration (-21.74dB) trials. We wish to know whether the Arc Scribing technique was significantly different from the other 3 unspecified techniques. Figure 73 shows the psychometric curves generated by the Jorgensen algorithm for the 4 tests administered to this subject together with the calculated errors in the thresholds. The circles indicate stimulus levels actually presented to the subject as specified by the QUEST algorithm.

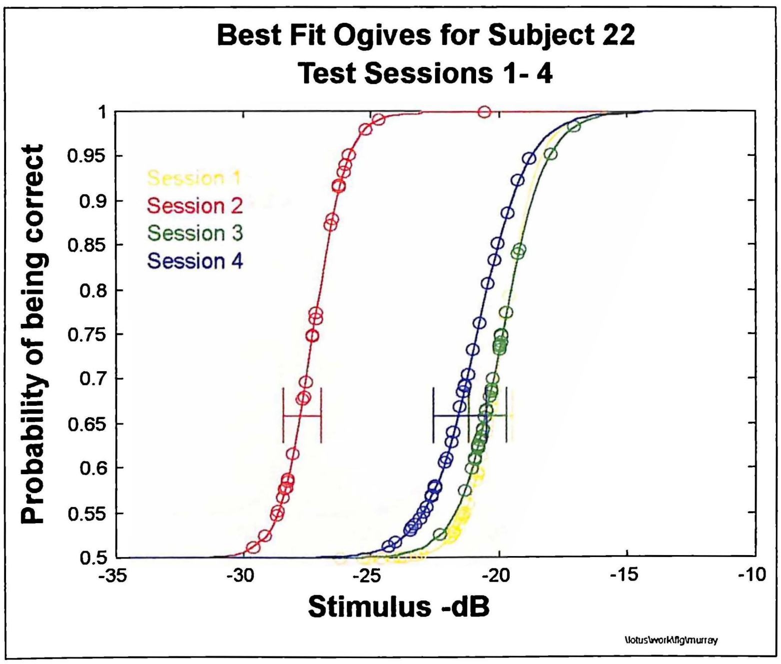


Figure 73

Figure 73 shows that the test sessions where the subject did not report using any special techniques all produced very similar psychometric curves. The test session where subject 22 reported using the Arc Scribing technique produced an ogive that was well separated from these other curves. There are a couple of interesting features to notice about Figure 73;

i) The Jorgensen algorithm produced larger estimates of the errors in the thresholds than were produced during our preliminary calculations in section 6.3, in line with the findings of McKee, Klein and Teller (1985)

ii) The QUEST algorithm specified stimuli close to the calculated thresholds but in general slightly biased to the left of the threshold - again in line with the findings of McKee, Klein and Teller (1985)

We can combine the 3 sessions without any reported technique to yield Figure 74 below.

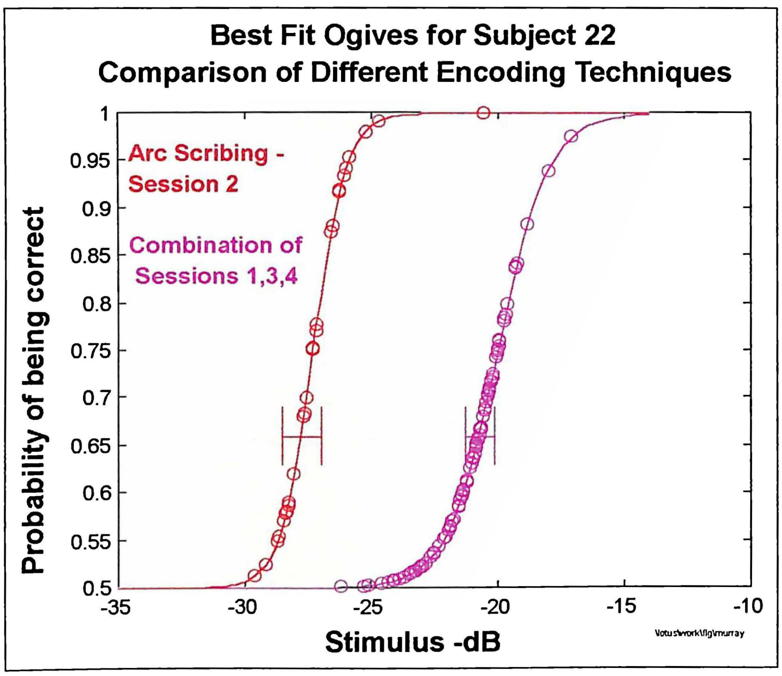


Figure 74

Figure 74 shows that combining sessions 1, 3, and 4 produced a single curve with a small error in the threshold estimate, indicating that the 3 sessions are likely to represent the same underlying performance. Of equal interest to us are the likelihood surfaces associated with the thresholds and gradients. Figure 75 shows such a surface for session 2 from Figure 73, while Figure 76 shows the corresponding contour plot.

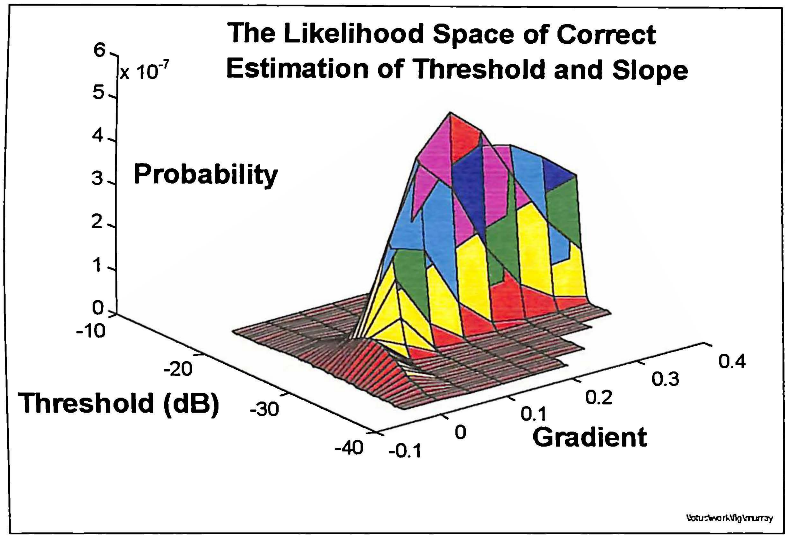


Figure 75

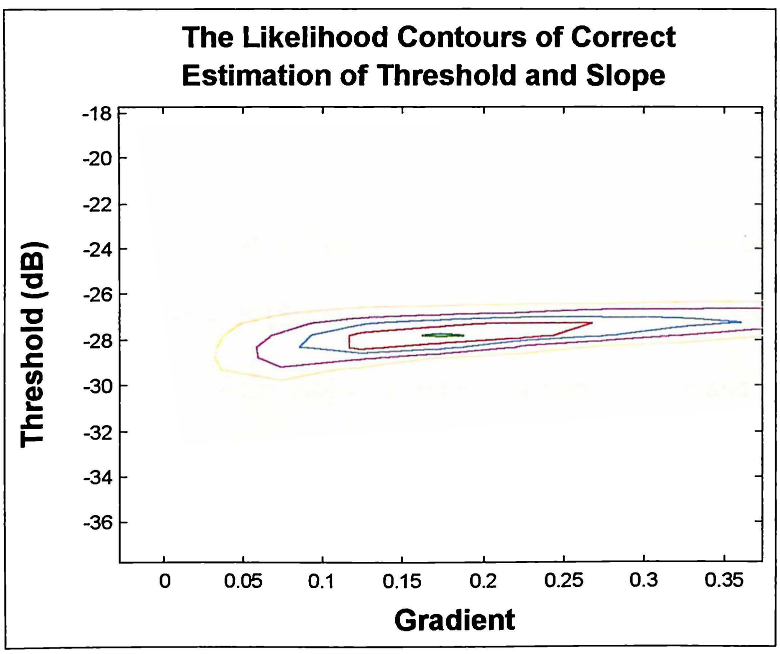


Figure 76

Figures 75 and 76 show that the threshold estimate is indeed well determined as indicated by the horizontal shape of the contours. However this same shape also means that the gradient of the psychometric curve is not well specified although this is not of prime importance for the current discussion and is a natural consequence of the way the QUEST algorithm attempts to place the stimuli close to the subject's threshold.

Figure 77 shows the likelihood surface in the parameter plane (a_i and b_i).

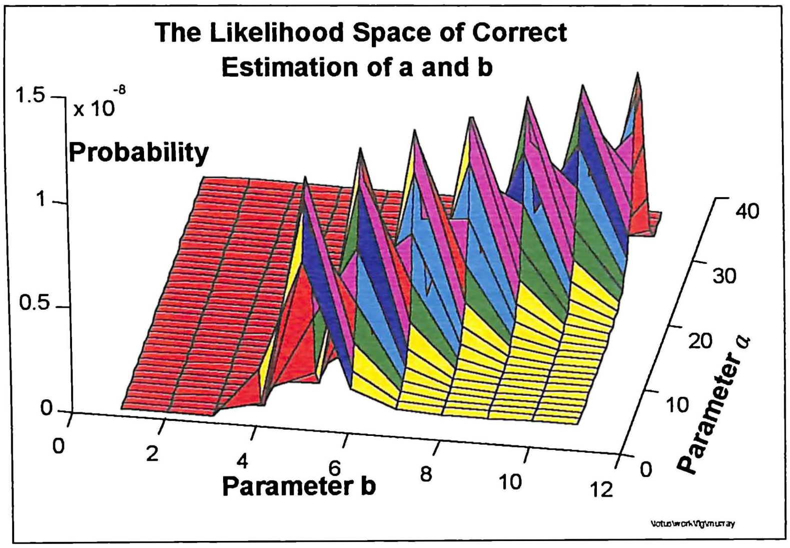


Figure 77

Figures 73 to 77 indicate that it is extremely likely that the different thresholds obtained by the two methods (Special Encoding vs no reported technique) for subject 22 were due to a change in the subject's underlying performance and not simply a statistical anomaly. The results for subject 22 were quite typical of the results obtained for the other subjects. Figure 78 shows the psychometric curves for subject 1 who used the musical encoding scheme discussed in section 6.6 while Figure 79 shows the psychometric curves for subject 16 who used the golf encoding scheme.

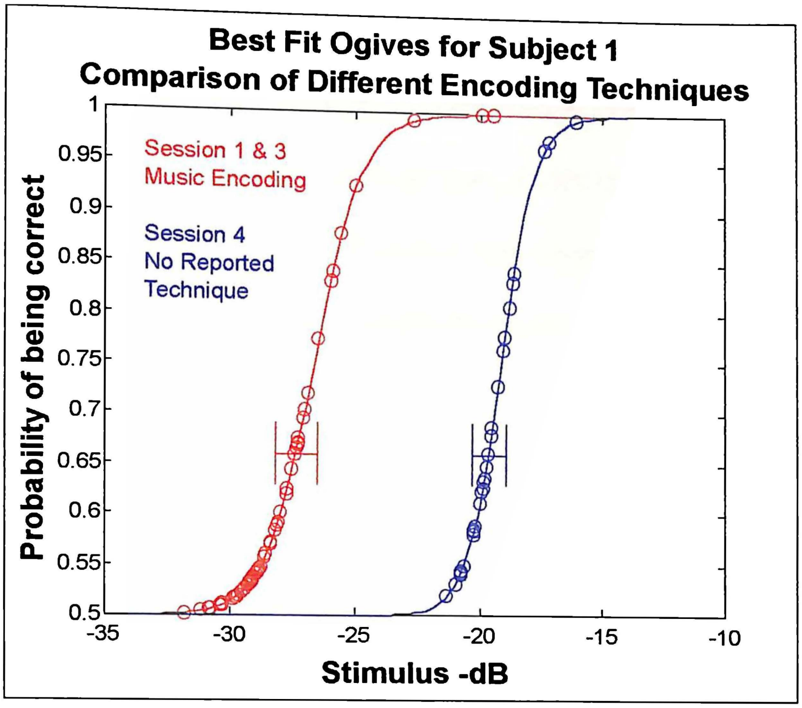


Figure 78

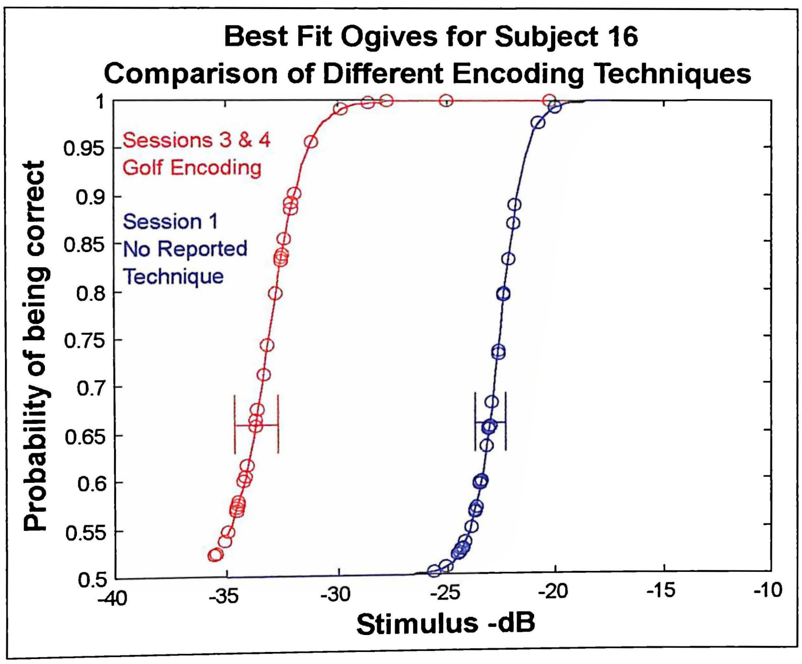


Figure 79

Therefore it appears very likely that vestibular discrimination thresholds can be strongly influenced by higher cognitive factors. If this is true then it means we may be able to substantially improve the vestibular discrimination performance of the VDEs which in turn may mean we can improve their vehicle assessment performance. This possibility was illustrated by the improvements shown by subjects 1 and 16 in the vestibular tests which were then reflected in their Lane Change results. This possibility has led us to develop a much more substantial vestibular research and training program for 1996.

6.8 Future Vestibular Research and Training- Motion Gymnasium

In 1995 subjects attended the course in groups of 6 and were simultaneously assigned to various modules in a random order. For 1996, subjects will be individually trained and tested over a 3 day period which will allow us to conduct formal testing on them in a controlled fashion and in the correct sequence over the entire 3 day period. Ford management also now appreciates the benefits of a formal research program and are more relaxed about a significant percentage of VDE contact time being utilised for that purpose. In addition, as the 1996 program will last for 16 weeks and as we will now have our own vestibular facility permanently located at the test track, we will be able to conduct long term tests by measuring the vestibular performance of each staff member every 3rd day over the full 16 week period. This will allow us to try repeated tests using the ABAB format to determine the exact significance of different interventions for each subject. We wish to identify individual differences as well as group trends. This is most important as almost invariably the psychological literature on vestibular performance relates to group means and not to individual differences. What we clearly need is an ability to reliably separate out individuals and determine the effects of training on each of them.

Our Motion Gymnasium will consist of a single axis Rotary platform whose motion can be precisely controlled by computer. Subjects will be exposed to a range of stimuli and trained using feedback techniques. Training will be as important as assessment. In some cases the subjects will be passively moved and simply have to report on the motion (as per the current research) while in other cases they will use a steering wheel to respond to the motion, for example by attempting to null the applied motion. The addition of the steering wheel will allow considerable extension to the investigations that we can undertake. We have already seen that tracking performance using visual information is quite different to that using vestibular or combined visual and vestibular information and that this difference is more profound as the frequency increases. It may be that what is most important in evaluating a vehicle is not just the actual motion of the vehicle but the relation of the motion to the inputs applied by the VDE. The subjects may be basing their discrimination during the Lane Change on the small instantaneous differences between the steering inputs and the motion outputs. Differences in measures such as phase delays and short term instantaneous response gains would be exactly the type of things that a high frequency response system like the vestibular system should be good at discriminating. Even though the overall motion of the vehicle during the Lane Change has a large low frequency component, there are also considerable short term high frequency components as the vehicle moves from one section of the Lane Change to another. Identifying and responding to these high frequency changes is exactly what is required from an elite racing driver like Jackie Stewart who performed exceptionally well in the Lane Change. As we shall see in the Attachment on Elite Race Driver training, Jackie Stewart was easily able to null out a Yaw instability in a racing vehicle whereas another elite racing driver in the same vehicle could not.

Thus what we aim to do is build up a profile for each subject in both the passive and active modes and determine their operating characteristics across a range of frequencies. We then wish to establish the effect of training on these measures. Are there some components that are amenable to training and are there others that are determined primarily by physiology? It could well be that for example, the low frequency memory components of motion can be improved with training while the high frequency response components are hard wired.

Another area that we will investigate is the effect of somatosensory input on supra-liminal discrimination. While the research of Walsh (1961) showed that subjects with spinal lesions who had no somatosensory input from their lower body performed identically to normal subjects in very low frequency absolute detection threshold experiments, this finding may not be applicable to motions that are of higher frequency or for those that are well above detection threshold as in the Lane Change manoeuvre. Just as the experiments of Meiry (1965) showed that the addition of vision produced no significant improvement in performance at low frequency but it did at high frequencies, so too we may find that this is the case with the somatosensory inputs. Racing drivers and pilots often refer to driving or flying 'by the seat of their pants'. This is more than just a turn of phrase. When Jackie Stewart is required to perform a critical evaluation he will often request that a special seat is fitted to the vehicle to help him detect the subtleties of motion. To test the influence of the somatic sense on vestibular performance we will use two techniques. In the first technique we will mount different chairs and determine their effect on performance.

We will use both a standard passenger vehicle bucket chair and a rally seat.

In the second technique we will introduce a masking vibration by the placement of a variable speed vibrator under the chair.

7 Theoretical Model of Attention and Vestibular Performance

In this section we develop a theoretical model which can account for improvements in Lane Change and vestibular test discrimination performance when attention is restricted to one only part of the motion waveform. We do this by drawing together experimental findings from the visual and auditory senses and extend these to the vestibular sense. While this section is not central to the main thesis developed later in section 9, it provides a backdrop to some of the issues involved in stimulus perception which turn out to be crucial in that section.

When we are presented with a stimulus, such as a visual pattern or a sound, we do not simply make direct use of the stimulus per se. Instead we process the stimulus to extract out its most important features to help us interpret the information. But analysing these sensory inputs to determine their most salient features may take longer than the stimulus is available to us, or at least is available to us in a constant enough state to be able to perform the analysis. (Lindsay and Norman, 1972).To overcome this problem, the Short Term Sensory Store (STSS) provides an accurate working copy of the stimulus for a short period of time which allows us to perform this feature extraction and analysis procedure. For example, a visual image remains for several tenths of a second after the visual input has been received. This means that it is possible to work on the sensory event for a duration of time which is longer than the event itself.

Not only does the STSS seem to retain a good copy of the sensory events that have occurred during the past few tenths of a second, but there is often more information stored in the STSS than can be extracted. This discrepancy between the amount of information held in the sensory system and the amount that can be used by later stages of analysis is very important. It implies some sort of limit on capacity at later stages of analysis, encoding or storage, a limit that is not shared by the sensory transduction stages themselves. The limitation shows up during the attempt to remember the material presented. To illustrate the nature of the STSS, let us consider the experiment conducted by Sperling (1959) on the visual STSS, shown below in Figure 80.

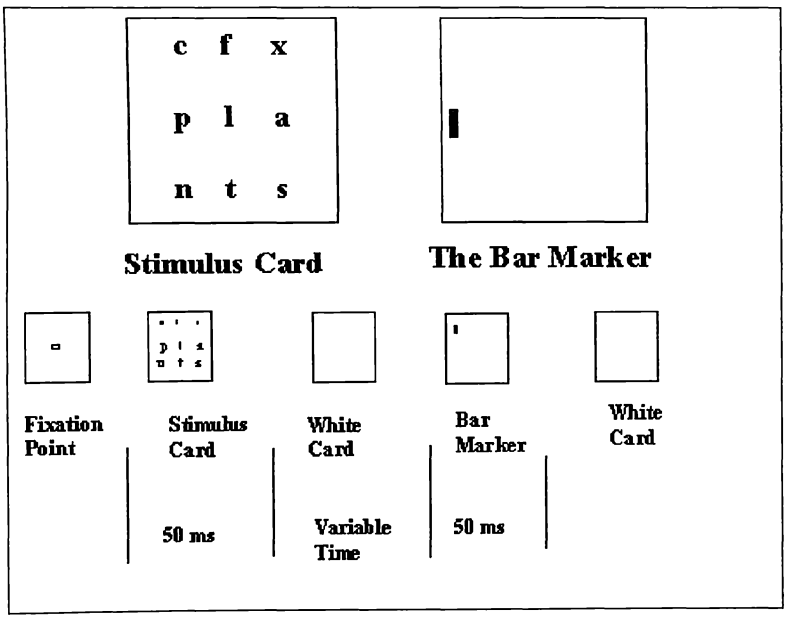


Figure 80 Schematic Outline of Sperling Experiment

(after Sperling, 1959)

The experiment is performed in two parts. In the first part, subjects fixate on a white card which has a cross at the centre. This is followed by the presentation for 50ms of a stimulus card which contains a number of letters. Following this, the subject is again presented with a blank white card during which he has to recall as many letters from the stimulus card as possible. Typically the subject manages only to record 4 or 5 letters. Even if the number of letters is increased or the exposure duration varied, the number of letters they report remains approximately the same. This is shown below in Figure 81.

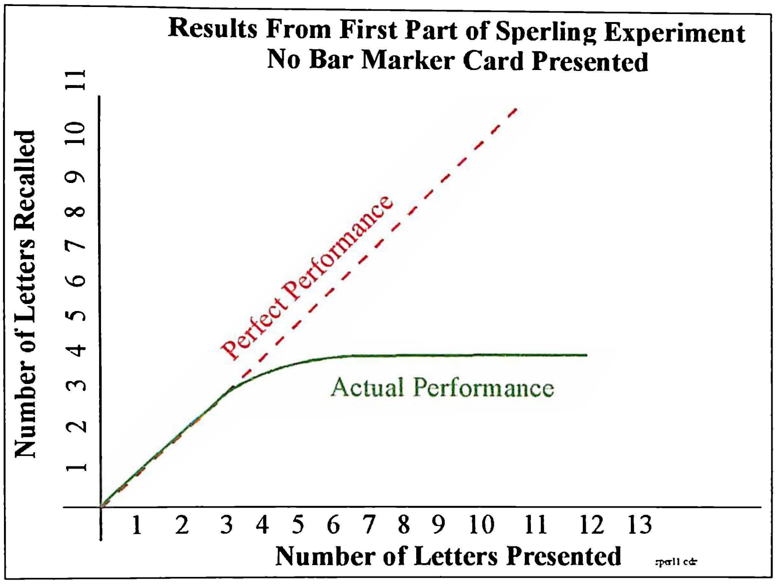


Figure 81 (Sperling, 1959)

In the second part of the experiment, the subject is again presented with the fixation card followed by the stimulus card for 50 ms and then the white card, but after this, the subject is presented with the bar marker card. The subject now has to report which letter was in the position beside the bar marker. The subject never knows which of the 9 letters is going to be marked until after the stimulus has been removed. Now if the subject can report on any randomly marked letter, then they must be able to 'see' all the nine letters at the time of the recall. They must be able to search for the letter that was in the position indicated by the marker and have that letter available. When the second part of the experiment is performed, the subject can almost always correctly identify the marked letter as shown in Figure 82.

This supports the hypothesis that subjects do have more information available to them than they can process, but that this information quickly fades.

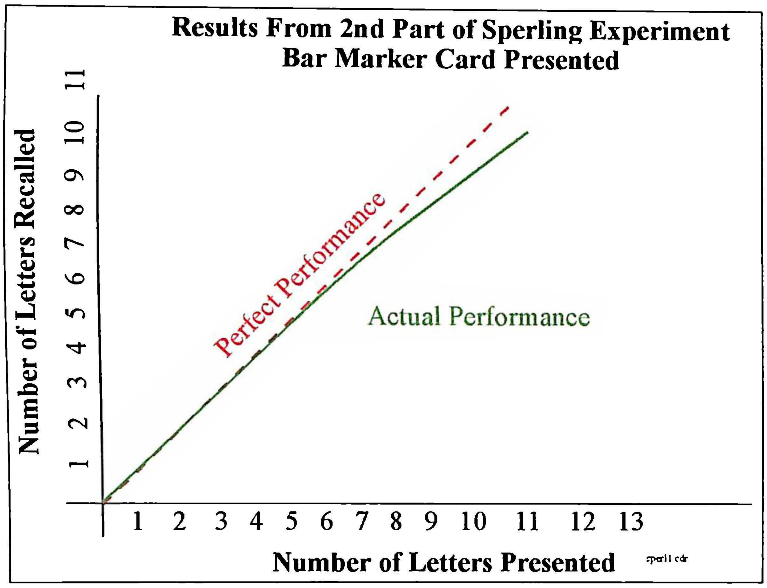


Figure 82 (Sperling, 1959)

In the first part of the experiment without the bar marker, all the letters were in the STSS, but by the time the subject had processed and recalled 3 or 4 of them, the others had faded away. In a variation of the experiment, subjects are asked to recall all the letters that occur in the same row as the marker. When this is done subjects can recall all the letters in the row indicating that all the letters must have been available to them.

Additional information can be obtained about the nature of the STSS by delaying the onset of the bar marker. By increasing the delay in presenting the bar marker, we can determine the duration of the STSS. Figure 83 below shows typical results obtained from such an experiment.

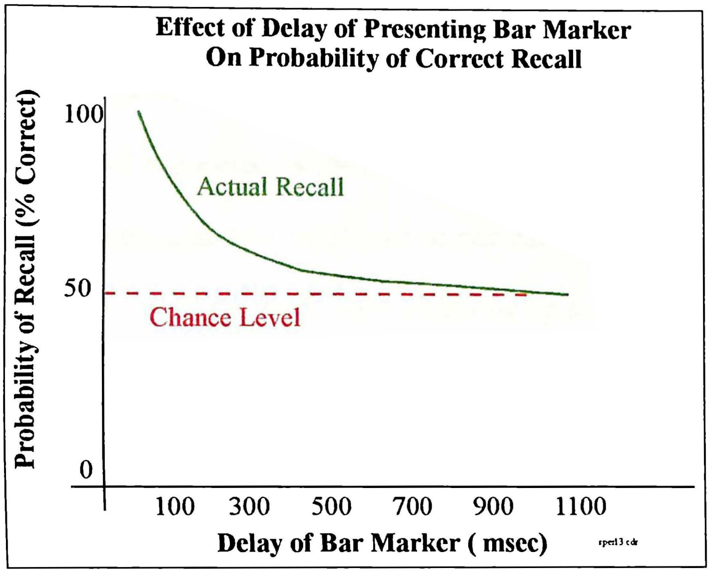


Figure 83

From Figure 83, it appears as if the STSS decays exponentially with a time constant of approximately 150ms so that after about 0.5 seconds, there is little of the image remaining.

Similarly evidence exists for an auditory STSS. Treisman (1964) used the repetition method to investigate the auditory STSS by presenting two identical coherent verbal messages to each ear of a subject. However, one of the messages was delayed with respect to the other. The subject's task was to shadow or repeat aloud the message received in one of the ears which was specified in advance by Treisman. Treisman's assumption was that shadowing the message in the designated ear was fully demanding of the subject's attention so that any memory for the message in the unattended ear must be based completely on pre-categorical sensory memory. That is the information available from the unattended ear will not have been processed to any higher level, will not have had any linguistic features extracted from it, and will only be available to the subject for comparison if it is indeed held in some form of unprocessed short term sensory store. This view that unattended semantic information is not generally available to a subject is well supported by the literature. For example Moray (1959) had subjects shadow words presented over earphones to one ear while common English words were presented to their other ear, each word occurring up to 35 times. When the experiment was over, the subjects' memory for the words presented to the unattended ear was tested. It was found that there was absolutely no memory for these words. Evidently, the attention required to do the shadowing task completely disrupted the subjects' ability to deal with the information presented to the other ear. Now the point of interest in Treisman's experiment was that she varied the time delay between the two messages and also whether the unattended message led or lagged the attended message. The behaviour of interest was whether the subject detected that the same message was being repeated in the unattended ear and how this detection depended on the two independent variables - delay time and lag or lead.

When the unattended channel led the attended channel, the delay had to be less than 1.5 seconds before the subject noticed that the two messages were the same. However, when the attended channel led the unattended channel, the delays could be extended to intervals greater than 4.5 seconds. The discrepancy between these two intervals makes the experiment a convincing one. When the attended channel leads the unattended channel, the information in the attended channel has been subjected to a full categorical analysis and should therefore be available in short term verbal memory, which is quite distinct from the Short Term Sensory Store we are discussing, for comparison with the message in the unattended ear. But when the message in the unattended ear leads the message in the attended ear, the only information available to the subject to compare with the current message should be that which has not been processed and transferred to the verbal short term memory, i.e. that which is contained in the auditory short term sensory store or echoic memory. Thus, Treisman's experiment supports the concept of an auditory short term sensory store of approximately 1 or 2 seconds. Figure 84 shows this schematically.

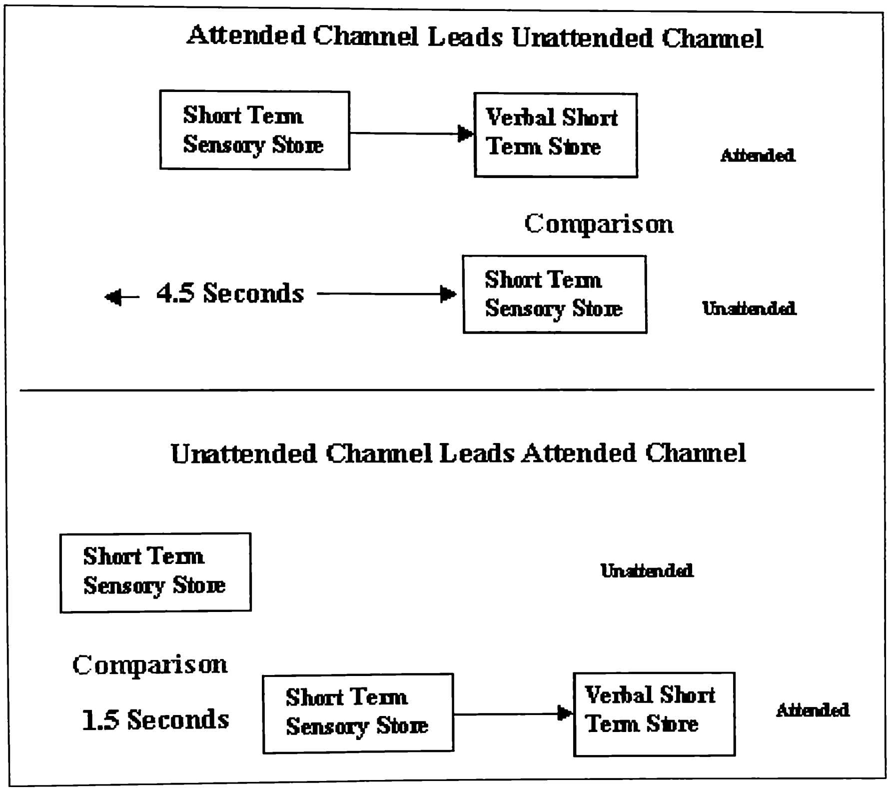


Figure 84 Schematic Outline of Treisman's Experiment on Auditory Short Term Sensory Store

(After Treisman, 1964)

There is a notable distinction between the Sperling and the Treisman experiments in that the different components of the information in the auditory experiments of Treisman are presented to the subjects sequentially whereas the different components in the Sperling experiments are presented all at once or in parallel. Our vestibular experiments more closely approximate the Treisman experiment as the motion information is presented to the subjects in sequential fashion. There is a single perception of motion in one axis which is varied with time.

Returning to the Treisman experiment, it is important to remember that the subject is required to determine whether the two messages are linguistically the same or different. This requires a comparison at the semantic level rather than at just the pure auditory level. This distinction is important because other experiments (Moray, 1959; Cherry, 1953) where unrelated messages were simultaneously presented to both ears (dichotic listening) have shown that for the unattended ear, the subject is subsequently able to report at the conclusion of the experiment from their long term memory;

i) Whether an unattended message was presented or not

ii) Whether the voice changed from male to female or vice versa

iii) Whether other unusual signals such as whistles etc. were presented

But the subject was not able to;

i) Report the contents of the unattended message

ii) What language the message was in

iii) Tell if the language changed during the course of the experiment

iv) Distinguish legitimate speech sounds from non linguistic sounds comprised of incorrectly ordered phonemes

These results indicate that attention is necessary to extract higher level features from the auditory data but that very coarse levels of information such as male vs female voice can still be extracted without attention.

We now look at the effect of detection and discrimination difficulty on processing time requirements. Chocolle (1940) has shown that as stimulus intensity is reduced towards a subject's threshold, the subject's reaction time required to decide if a stimulus is present or not is increased. Figure 85 shows data obtained in a classical auditory detection experiment where the subject simply has to press a key if they detect the presence of a 1000Hz tone.

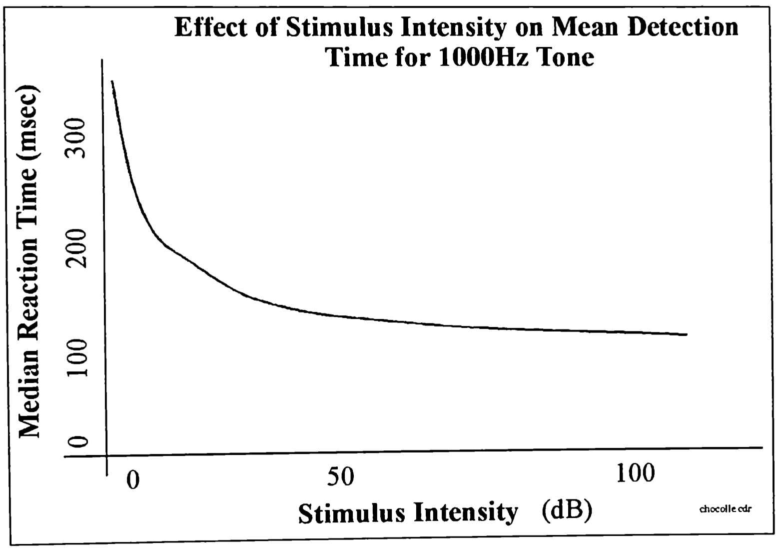


Figure 85 (From Chocolle, 1940)

Similar results have been obtained in discrimination tests. Welford (1980) found that when subjects had to discriminate between two objects, the larger the difference the shorter the reaction time and hence by inference the shorter the processing time. Shallice and Vickers (1964) required subjects to discriminate two lines on the basis of their length. The subjects were simultaneously presented with two lines on a card and simply had to sort the cards into piles on the basis of which of the lines was the shortest of the two. Again, the greater the difference in line length, the quicker the subjects completed the task.

Thus we can conclude that the more complicated the feature analysis, extraction and encoding process required from the raw sensory stimulus, the longer it takes to complete and the more vital is the allocation of attention to its success.

But how exactly does altering attention affect a detection or discrimination task? Posner (1978, 1980) presented subjects with cue fixation cards prior to presenting them with a target. The subjects simply had to press a button as soon as they detected the presence of the target which consisted of a square box which could occur on either side of the stimulus card. There were 3 experimental conditions called neutral, valid and invalid. In the neutral condition, the cue card consisted of a central fixation cross and the target square would occur randomly at 50% probability either to the left or to the right of where the fixation cross occurred. In the valid condition, the cue card consisted of an arrow centrally located which pointed to one side of the cue card and this was followed by the stimulus card which had the target square on the side to which the arrow was pointing. In the invalid condition, the target square occurred on the side opposite to which the arrow was pointing. The invalid condition occurred on 20% of the trials where the arrow cue appeared and the valid condition occurred on the remaining 80%. This is shown schematically below in Figure 86.

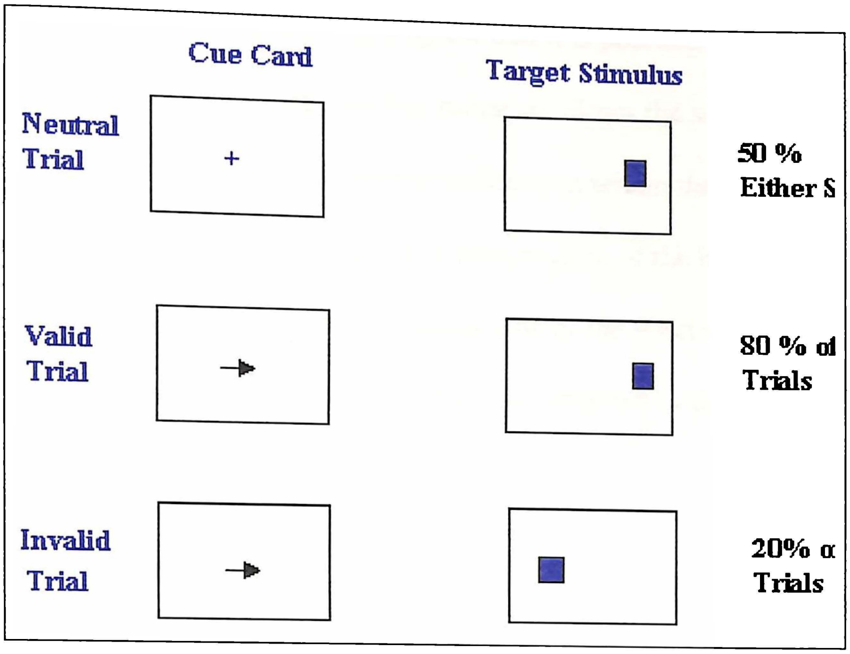


Figure 86 Schematic Outline of Posner's Experiment on Attention Direction

(Posner, 1978)

In all cases the subjects were not allowed to move their eyes around the visual field but were required to keep them focused at the centre of the visual field. The results of this experiment are shown below in Figure 87.

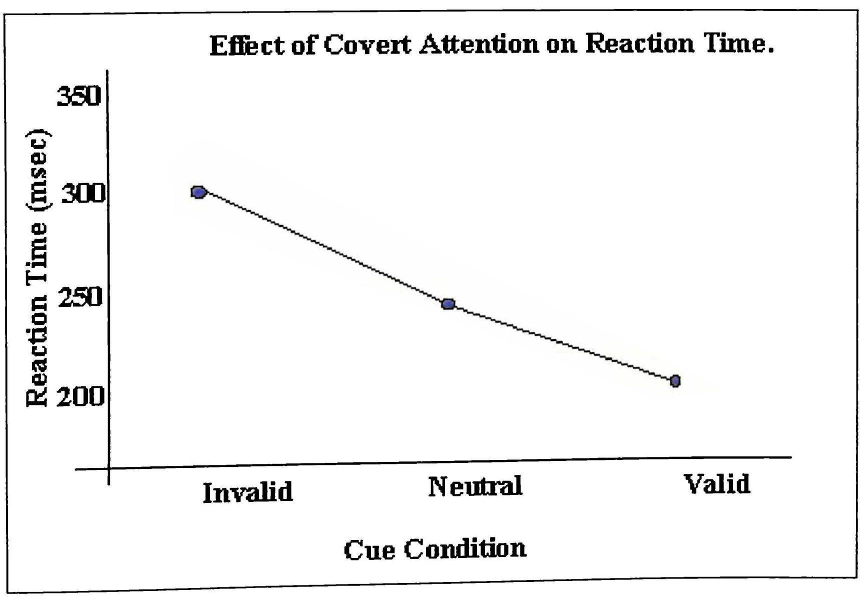


Figure 87 (Posner, 1978)

The results of the Posner experiment suggest that it is possible to mentally direct one's attention within the visual STSS and that doing so allows the subject to process the information at that point faster. This mental direction within the STSS supports the Sperling experiment where the subsequent presentation of the bar marker card served to direct the subject's attention to a particular row of the 9 letter stimulus matrix which was now no longer physically present, but which remained available to the subject only in their short term sensory store.

Stelmach and Herdman (1991) conducted an investigation which dramatically showed the effect of attention on processing speed. In this investigation subjects were required to fixate on a central square and were then presented with display dots which occurred on either side of their fixation point. The left and right dots could be presented at the same time or they could be presented sequentially with different stimulus onset asynchronies (SOA). The subjects had to report using the 2 alternate forced choice procedure whether the left pattern of dots or the right pattern of dots appeared first. The PEST algorithm (Taylor and Creelman, 1967) was used to adjust the SOA to find the asynchrony at which the subjects could not differentiate the order of the stimuli. The key part of the experiment was that while the subjects were required to keep their eyes fixed on the central square they were also directed to shift their <u>attention</u> either to the right or to the left. In some trials the subjects attended to the left and in other trials the subjects attended to the right. The subjects' eye positions were monitored to ensure that they kept looking at the central square regardless of where their attention was directed. Stelmach and Herdman found that dots on the attended side were judged to appear simultaneously with the unattended dots only when the unattended dots were presented first with an SOA of approximately 40 ms.

This is shown below in Figure 88.

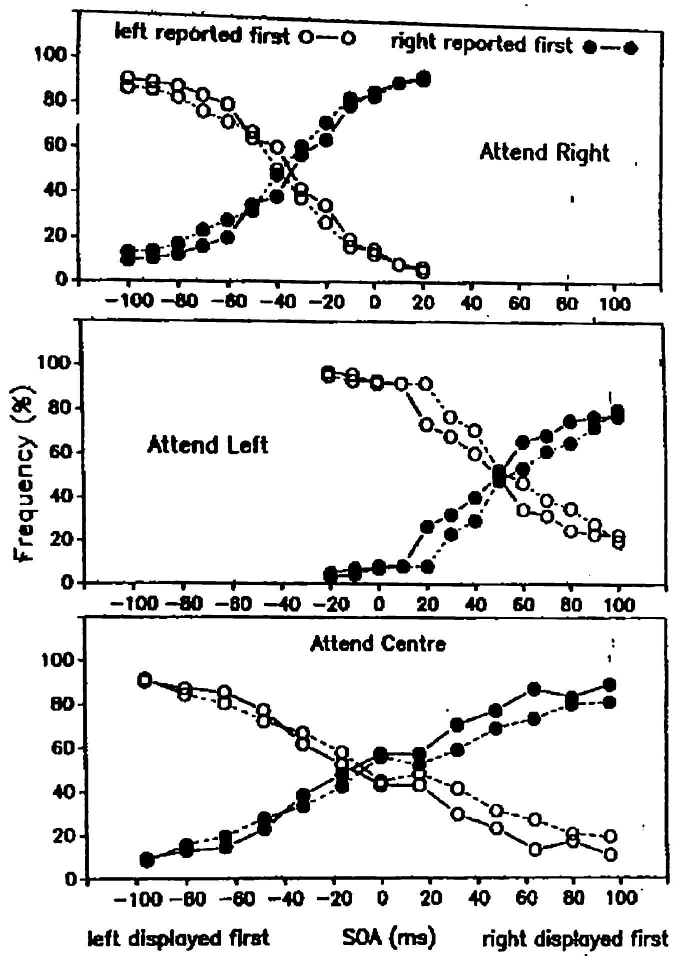


Figure 4. Frequency of left-first and right-first responses (Experiment 3) across SOAs when eye movements were monitored (dashed lines) and not monitored (solid lines).

Figure 88 Effect of Orienting Attention on the Perception of Simultaneity for Asynchronously Presented Stimuli

(Stelmach and Herdman, 1991)

Thus stimuli which are equally separated from the point of fixation within the visual field are detected at different speeds depending on where in that field the subject has directed their attention. This experiment is particularly convincing because the effect of attention on detection speed is not dependent on the amount of time the subject takes to respond to the forced choice. There is thus no possibility of confusing other cognitive factors such as detection time with detection speed.

The above investigations might give the impression that the allocation of visual attention is a similar to a spotlight where there is a fixed point of focus. Eriksen and Yei (1985) showed that this is not the case. In their experiments they showed that attention could be spread out over a wide visual field to provide lower resolving and processing power or could be concentrated on a smaller region within the visual field to provide a greater level of processing within that restricted region. In their experiments attention was thus focused in a manner analogous to a zoom lens. Driver and Baylis (1989) showed that attention need not even be allocated contiguously over space. They showed that attention could be directed to perceptual groups whose components were spatially dispersed but which share common features, in their case motion. Subjects could be directed to attend to a number of discrete objects that moved at a specific speed compared to other objects that moved at different speeds, despite the fact that these common speed objects were widely separated in the visual field. These widely spaced but attended objects could be processed better than unattended objects.

We can now combine these experimental results into a single hypothesis which explains why restricting the subjects' attentional focus during the vestibular and the Lane Change Discrimination experiments improved their performance. Figure 89 shows a schematic representation of the flow of vestibular information perceived by the subject during a single trial of the vestibular discrimination experiment.

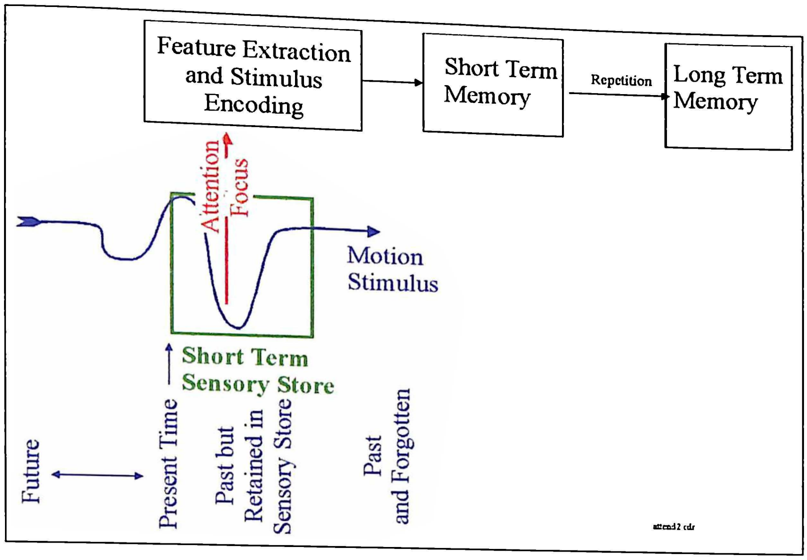


Figure 89 Schematic of Vestibular Information Processing

The green rectangle schematically represents the time history of the subject's Vestibular Short Term Sensory Store. The left hand side of the STSS rectangle corresponds to the present while moving further to right in the STSS indicates information which has been held in the STSS for longer periods of time and hence has become more degraded. The blue waveform indicates the motion that the subject will or has experienced. In this model, the subject is able to shift their focus of attention around within the STSS in much the same manner as postulated by Posner and Treisman. This is indicated by the vertical red line. As per the previous experiments, it is assumed that the component of the vestibular signal that is receiving the greatest attention will be processed more accurately, faster and to a deeper level than other components of the sensory signal.

We can now diagrammatically depict the two different strategies used by the subjects;

i) When the subject concentrates all their attention on the peak waveform either as it is being presented or afterwards as it fades from the STSS

ii) When the subject tries to assimilate the entire waveform by focusing their attention on the currently experienced stimulus

These two strategies are shown schematically below in Figure 90. On the left hand side we see that the subject's attention remains focused on the peak amplitude while on the right hand diagram we see that the subject's attention remains focused on the current incoming stimulus, which as we have seen is the position of strongest perception before the stimulus is degraded by decay.

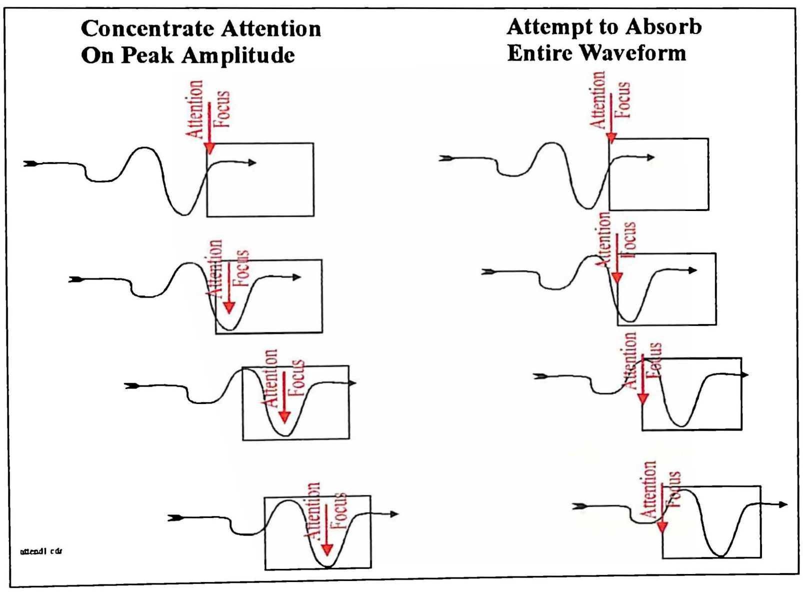


Figure 90 Schematic Representation of Applying Different Strategies to Vestibular Information Encoding

Now if it is true that;

i) The vestibular stimulus is indeed held in a short term sensory store for a brief time and

ii) Focusing attention on that experience while it persists in the STSS allows additional information to be extracted from that part of the stimulus and encoded

then restricting the subject's focus of attention to just the peak stimulus experience should allow the subject to more accurately analyse and encode the magnitude of the peak of the stimulus at the expense of lesser analysis of the rest of the stimulus. Conversely, if the subject always attends to the current incoming information then there may be some interference of the current attended signal with the historically attended signal. This is similar to the Sperling experiment. For a brief period of time much information is available to the subject but the act of attending to a subset of that information, processing it and recalling it means that there was insufficient time to do so for the remaining information.

The important point to remember is that in the attenuated motion, the amplitude of the entire waveform is simply scaled by a constant factor (a fraction less than 1). This means that the size of the difference between the Base and the Modified motions (dS) at any point along their stimulus presentation is proportional to the magnitude of the stimulus (S) at that point,

that is, $dS = k.S$

But this is identical to Weber's law except in that case dS is the minimum discriminable difference in stimulus at a stimulus level S, and k is the Weber fraction.

Now, the Weber fraction is not a constant but is always initially decreasing and convex (Boring, Langfield and Weld, 1948; Laming, 1983; Luce and Green, 1974; Green, 1978; Falmagne 1985). It is only ever increasing at the upper end of a sensory mechanism's range (Holway and Pratt, 1936). Because the vestibular signal in our test ranges from zero to approximately 30 times threshold (Clark and Stewart, 1974; Fenessay, 1975; Benson, Hutt and Brown, 1989) but less than 20 times tolerance (Benson, 1990), this means that the waveform is most discriminable at the positions of peak velocity or acceleration (depending on which the subject is using) which is where the Weber fraction is lowest. That is there exists a section of the vestibular platform motion waveform where more useful information is available to the subject (i.e., the signal to noise ratio is highest) than other sections of the waveform. The same is true for the Lane Change where the biggest difference in vehicle behaviour (as measured by the telemetry) occurs at the exit to the transition section. This presentation of high quality information is preceded by and followed by much longer periods of lower quality information in both the Lane Change and the vestibular tests.

By concentrating their attention on the peak event (or in the case of the Lane Change, the most significant event), subjects will process the short duration higher quality stimulus better at the expense of less processing of the bulk remaining lower quality stimulus. This strategy will yield better discrimination if the gain in information from the high quality stimulus region exceeds the loss of information from the remaining low quality stimulus region. That this appears to occur in both the Lane Change and in the vestibular tests which indicates that there is a significant amount of processing required to extract the optimal amount of information from the high quality stimulus region.

The important point to realise with respect to the Lane Change test was that there was a well defined region which contained the most information and further that the subjects were <u>already aware of this and concentrating on</u> it prior to training. It wasn't as if we were directing their attention to a region that they had been previously neglecting. Rather, we were telling them to <u>continue</u> concentrating on that region even after it had passed. When the subjects did this they reported that they had a 'much clearer image' of the motion. They reported that they could detect subtleties that had previously 'passed them by'. It wasn't that the attention focus improved their memory for the motion but rather they appeared to be able to 'feel' much more to put into the memory in the first place. Our subjects reported that they were able to use these additional subtleties to improve their discrimination performance.

This reminds us of the key difference that we first noticed between Jackie Stewart and the other test drivers prior to conducting these programs (see section 6.1). After each run Jackie seemed to be able to 'replay' the most important parts of the motion in much greater detail than the other subjects. When the other subjects were shown the telemetry plots after their runs and the differences in the vehicle's behaviour at the exit of the transition section were pointed out to them, they usually reported that they 'never felt that'.

The preceding discussion raises the question whether there is historical evidence that discrimination thresholds are fixed entities or whether they can be affected by learning or practice. As far back as 1858 Volkman tested tactile discrimination performance. Specifically Volkman measured tactile acuity which is the smallest distance between two points depressed on the skin that are judged as separate rather than as single points. Volkman found that tactile acuity could be halved with practice. Furthermore, this improvement could be transferred to other areas of the skin that had not previously been tested.

In a replication of this experiment, Dressler found in 1894 that the two point threshold decreased from 21mm to 4.1 mm after 2000 trials over a 4 week period using the inner arm. Further, when training was confined to one arm but the other arm was tested before and after that training, it was found that the subjects produced near identical improvements in performance on the untrained arm, i.e. 21 mm before training and 5 mm on the untrained arm. This transference of improvement from one area to another is evidence of the effect of higher cognitive factors on discrimination performance.

In 1940 McFadden demonstrated improvements in visual acuity with practice by presenting subjects with cards containing a pair of black lines. The subjects had to report whether the lines were single or double. With practice the subjects could discriminate the two lines at increasing distances and hence at reduced visual angle. The above two examples however could be criticised using signal detection theory because there is the possibility that the improvements were due the subjects' change in discrimination criteria rather than in terms of their improvement in absolute sensitivity. To overcome these objections, Bjorkman and Ottander (1959) used the 2 alternate forced choice method to determine subjects' discriminability of lifted weights. They also instructed the subjects exactly how to lift the weights so that there was no difference in lifting techniques between successive trials. Their experiment showed that the difference limen dL was reduced by a factor of 2 over a period of 5 days practice.

The subjects were not given any feedback as to whether they were correct or not in the discriminations.

Baker and Osgood (1954) investigated discrimination of sound frequency and in particular investigated the effects of different types of training on the reduction in dL. They divided their subjects into 4 experimental groups after each subject had received an initial pre-test measure of frequency discrimination. The 4 groups were;

i) No training

ii) Training only on the test series of sounds

iii) Training first on an easy series of sounds followed by training on the test series of sounds

iv) Training on the easy series followed by training which gradually approached the test series in small increments of increasing difficulty

All subjects in groups b), c) and d) received the same amount of corrected practice trials. Baker and Osgood found that only subjects in group d) achieved significant improvement. The conclusion reached by them was that gradually increasing the difficulty of the discrimination helps the subject to define some hitherto unnoticed aspect of the difference relation which can then be used to serve as a distinctive or criterial feature for the judgement. Once identified the subject can focus their attention on this criteria.

A similar experiment was performed by Heimer and Tatz (1966) who presented pairs of recorded tones played against a white background noise. The subjects had to indicate whether the first tone was lower or higher than the second tone. The subjects were given 5 days of corrected practice following the pre-test. Again all subjects showed a significant improvement in their discrimination performance.

Thus there is historical support going back over a century for the idea that discrimination thresholds are not fixed entities, but can be influenced by training procedures or by different strategies.

8 Lane Change Discrimination Training

8.1 Introduction and Background

This section contains the main Lane Change research investigation conducted with the Ford VDEs and is the culmination of the preliminary research investigations. This research was conducted during May - November 1995.

The purpose of this investigation was to;

i) Establish the practised but untrained vehicle discrimination performance of the VDEs

ii) Determine if training procedures could improve discrimination performance

iii) Identify which training procedures (if any) provided the biggest improvement in performance

iv) Attempt to identify the factors limiting Lane Change discrimination performance

v) Determine if there was any correlation between Lane Change discrimination performance (either before or after training) and Vestibular test results (section 6)

iv) Determine if there was any correlation between Lane Change discrimination performance (either before or after training) and the RAF Fighter Pilot Attributes Selection test scores.

8.2 Method

8.2.1 Subjects

24 Professional VDEs selected by Ford management and comprised of 22 males and 2 females. The average age was 32 and most had at least a bachelors degree with 4 years postgraduate training and a minimum of 5 years VDE experience (std deviations not available as this data was supplied by Ford).

8.2.2 Vehicle

A 1995 Ford Explorer was instrumented by the fitment of a Cranfield RaceCorder data logger (Cranfield Institute of Technology) which has 12 bit resolution, 2 MB of RAM and 40 allocatable channels, 7 of which are digital. The logger was turned on and off automatically by infrared triggers which were positioned at the beginning and end of the Lane Change course and sampled at 50 Hz.

The following telemetry sensors were fitted to the Explorer;

Lateral Accelerometer - Schaevitz +/- 5 g Servo

Longitudinal Accelerometer - Schaevitz +/- 5 g Servo

Yaw Rate Gyro - Systron Donner Solid State

Throttle Position Sensor (L.V.D.T)

Steering Position Sensor (L.V.D.T)

Brake Position Sensor (L.V.D.T)

Brake Pedal Pressure Sensor (Strain Gauged)

Brake Line Pressure Sensor (Fluid PSI)

4 Suspension Travel Position Sensor (L.V.D.T)

Wheel Velocity Sensor (100 pulses per revolution)

The vehicle behaviour was modified by altering the front tyre pressures. The Base condition was defined as front tyres = 36 psi. The Modified settings ranged from 20 psi to 34 psi in 2 psi increments. Throughout all trials, the rear tyre pressures were set to 34 psi. Tyre pressures were checked after each run using precision tyre pressure gauges. Each subject used a new set of tyres which had previously been scrubbed in.

8.2.3 Analytical Techniques

Thirty six metrics were developed from this telemetry data which described in numerical format the overall characteristics of each run for each driver. Collapsing the telemetry data down to these descriptive metrics allowed statistical calculations to be performed on the drivers and allowed assessment of the efficacy of the different training techniques used. The metrics generated were of 3 types;

i) **Test Procedure** - Driver Number, Tyre Pressure Used , Type of Training, Driver Response etc.

ii) **Section Metrics** - Metrics summarising the Entry / Transition / Exit sections

iii) **Global Metrics** - Metrics which were dependent on the entire run

These metrics are listed below together with their metric number;

Test Procedure Metrics

1 = driver number

2 = run number

3 = tyre pressure

4 = confidence -ve if wrong

5= type of training session 0 = practice 1+ = training session number

5 = 1st discrimination test 6 = 2nd discrimination test.

Section Metrics

Entry	Trans	Exit	Metric
6	14	22	Steer Integral
7	15	23	Latacc Integral
8	16	24	Yaw Integral
9	17	25	Peak Steer
10	18	26	Peak Latacc
11	19	27	Peak Yaw
12	20	28	Steer Start Pos
13	21	29	Steer End Pos

Global Metrics

30 = Steer Responsiveness

31= Mean Transition Speed

32= Mean Trans Throttle

33= Latacc Hysteresis

34 = Yaw Hysteresis

35 = Peak Latacc / Steer

36 = Peak Yaw / Peak Steer

8.2.4 Test Track

Testing was conducted at the RAF Upper Heyford Airfield Base in England. The Lane Change was performed on the main straightaway. The Lane Change configuration used was identical to that used in the Meremere investigation and the Florida 1993 investigation as depicted in Figure 25 above.

8.2.5 Experimental Procedure

a) Instructions

Subjects were given a written manual prior to commencement of the course which contained full instructions on the Lane Change module including vehicle modifications, track layout, testing procedures and how they would be measured. On the first morning of each course before any testing commenced, all the subjects met together in a conference room and were verbally briefed by the experimenter on the Lane Change module. Subjects had a chance to ask any questions about the test procedure during the briefing. Finally, each subject was briefed personally on the procedure by the Experimenter immediately prior to their test. These final instructions were administered with the driver sitting in the driver's seat of the vehicle at the start of the Lane Change track so that each aspect of the procedure could be identified by the subject.

At the completion of the Lane Change module subjects were required to complete a written questionnaire where they rated the quality of various aspects of the module on a scale of 0 -10. All subjects rated the instructions at level 10.

b) Practice

Subjects were first given 20 trials of practice at driving the vehicle in the Base condition. They were instructed to ensure the vehicle maintained a constant speed of 45 m.p.h. throughout the Lane Change.

They were also instructed that they must use this time to;

i) Become familiar with how the vehicle behaved in the Base condition so that they could determine any changes induced by the Modified condition

ii) Develop a style that was very consistent and repeatable from trial to trial

iii) Develop a style which they thought would maximise the differences likely to occur in the vehicle's behaviour when the front tyre pressures were reduced

At the end of the 20 trial practice session, the subjects were given 2 sets of practice discrimination trials where the Modified pressures were set at the easiest level, 20 psi. The subjects were advised whether their answers were correct or incorrect at the completion of these 2 pairs of trials.

c) Discrimination Testing Procedure

Subjects were tested using the staircase algorithm (Cornsweet, 1962; Levitt, 1970) with the initial Modified stimulus set at the easiest level of 20 psi. Tyre pressures were adjusted in increments of 2 psi to a maximum of 34 psi for the most difficult Modified test. The test proceeded until the subject's discrimination threshold was found (using the criteria of 3 reversals). In some cases the test procedure was terminated if the subject failed to score any discriminations correctly or if they scored consistently at chance level on the easiest setting. In addition to identifying which was the Base and which was the Modified runs, subjects were required to rate their confidence of being correct on the previously established 0 - 10 scale.

d) Training

Following the establishment of each subject's practised but untrained discrimination threshold, each VDE underwent a training session lasting approximately 2 hours. This training included practical demonstrations of driving techniques, instruction on the psychology of sensory perception, motor control processes and how this applied to the

Lane Change procedure, telemetry feedback of their driving style and systematic sensitisation to dynamic stimuli.

e) Post Training Discrimination Threshold Testing

After each subject had received their individual training, they were again tested using the blind Discrimination Test procedure to determine a second measure of their discrimination threshold to see if the training procedures had improved their performance. This second Discrimination Test occurred at least 1 day after their final training session.

8.3 Results

8.3.1 Effect of Altering tyre Pressures on Vehicle Behaviour

Prior to analysing the performance of the drivers it is necessary to determine the effect of altering the front tyre pressures on the dynamic behaviour of the vehicle. A correlation analysis was conducted to identify which of the 36 metrics correlated most highly with changes to the front tyre pressures. This analysis revealed the following correlation coefficients in order;

Transition Steer Integral - metric 14	0.64
Peak Yaw / Peak Steer - metric 36	0.56
Latacc Hysteresis - metric 33	0.52
Yaw hysteresis - metric 34	0.42
Peak Latacc / Steer - metric 35	0.41

Table 5 Correlations of Metrics with Tyre Pressure Adjustments

All the above correlation coefficients were significant ($p < 0.001$) and were generated by finding the mean normalised metrics for each driver during both Discrimination tests at each tyre pressure setting to establish a single correlation for all drivers as a group. These results confirm that the Ford Explorer used in the Upper Heyford test behaved in a similar fashion to the Falcon XR-8 used during the Meremere investigation into enhancements to the Lane Change procedure. In particular, in both cases the Transition Steer integral was most highly correlated to the tyre pressure changes. This indicates that the test methodology of adjusting front tyre pressures appears to be transferable across completely different vehicle platforms which was one of our objectives in developing the enhancements at Meremere. (The Explorer is a light truck while the XR-8 is a sport sedan).

Figure 91 shows the magnitude of the effect of altering front tyre pressures on the Transition Steer integral and was constructed using the same technique as that used to construct Figure 44 in the Florida Pilot Study and Figure 47 in the Meremere study.

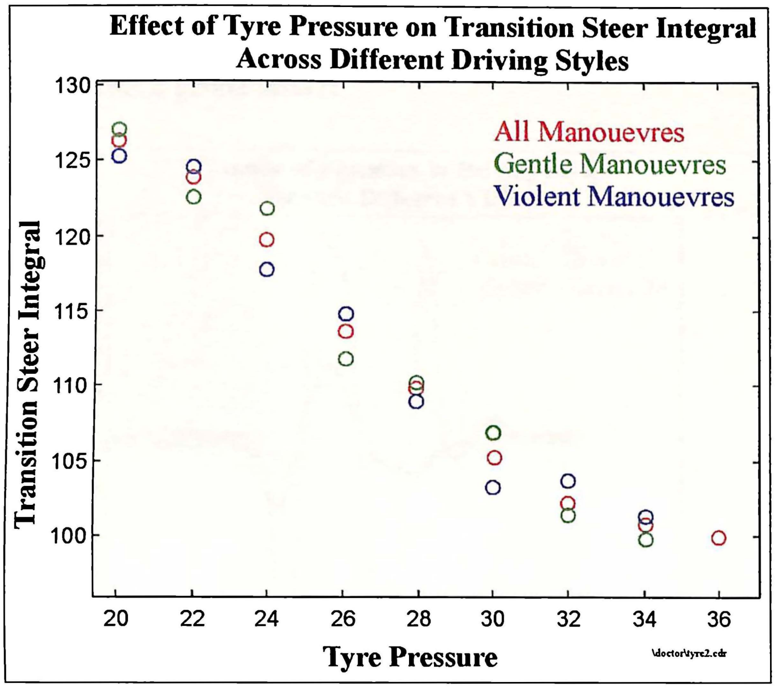


Figure 91

In Figure 91 we have plotted the average result for all the drivers in red to indicate the general trend. This shows that as the front tyre pressures are increased, the vehicle becomes more responsive to the steering input and the Transition Steer integral falls correspondingly. Initially the increase in steer responsiveness from the softest setting is linear but above approximately 28 psi the tyres become 'saturated' and further increases in tyre pressure produce smaller increases in responsiveness.

To determine whether this effect was robust across different driving styles we ranked the drivers in order of how violently they performed the Lane Change. The 25% most violent drivers, determined by their Peak Entry Steer Angle, are plotted in blue while the 25% most gentle drivers are plotted in green. Figure 91 shows that altering the tyre pressures produced equivalent changes in vehicle behaviour for both sets of drivers.

Figures 92 and 93 show the magnitude of the differences in driving style between the most severe and most gentle drivers.

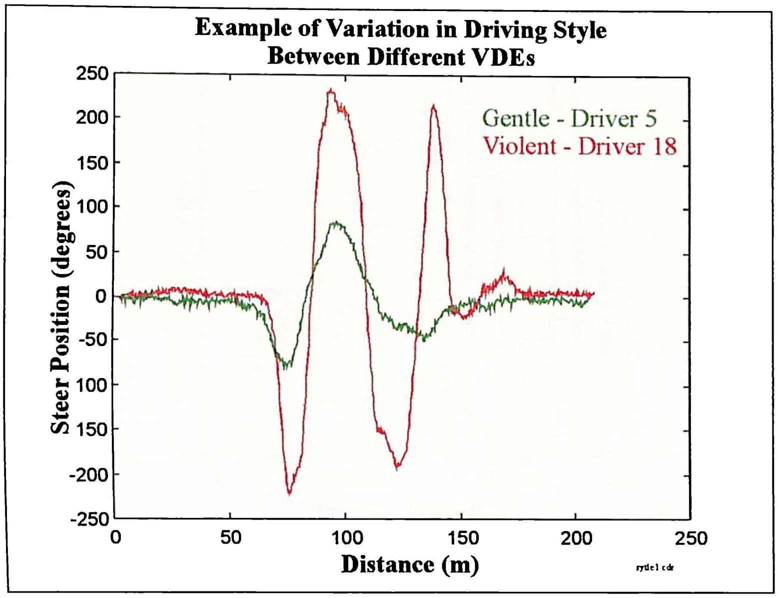


Figure 92

In Figure 92 we see that one of the drivers used approximately 3 times the Steer Angle to complete the Lane Change as the other driver. Different steer inputs will produce different transition steer integrals. Figure 93 shows the mean Transition Steer integrals for all 24 drivers during their 2nd Discrimination tests and shows that they differed by a factor of greater than 8!

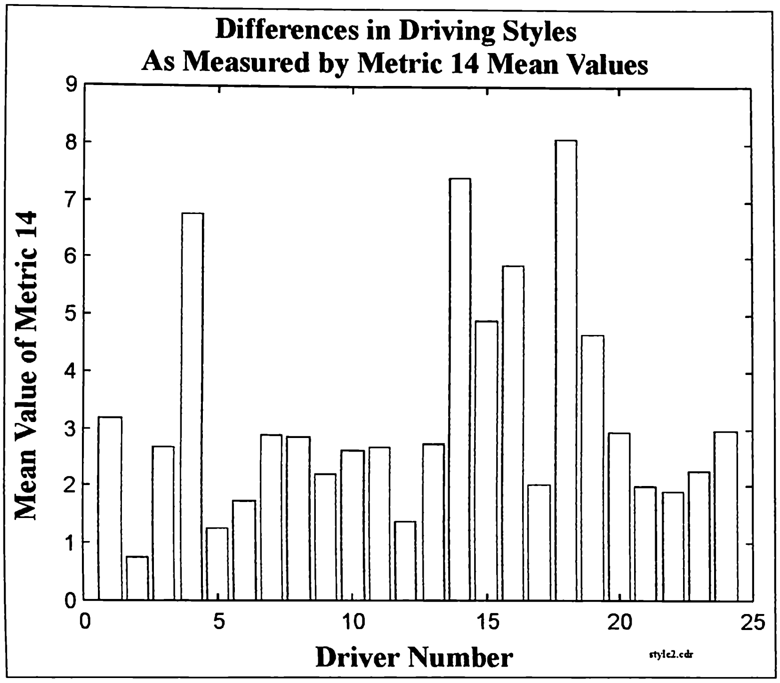


Figure 93

Thus changes to the dynamic behaviour of the vehicle caused by altering the front tyre pressures appeared to be robust across a wide range of driving styles.

Figures 94 to 96 show the effect that changes to the tyre pressures had on the secondary vehicle behaviours. If the vehicle behaved in a consistent fashion then reducing the Transition Steer integral should lead to a reduction in peak Yaw Rate which in turn should make the vehicle behave more linearly and hence reduce Yaw hysteresis. Figures 94 to 96 show that this was indeed the case.

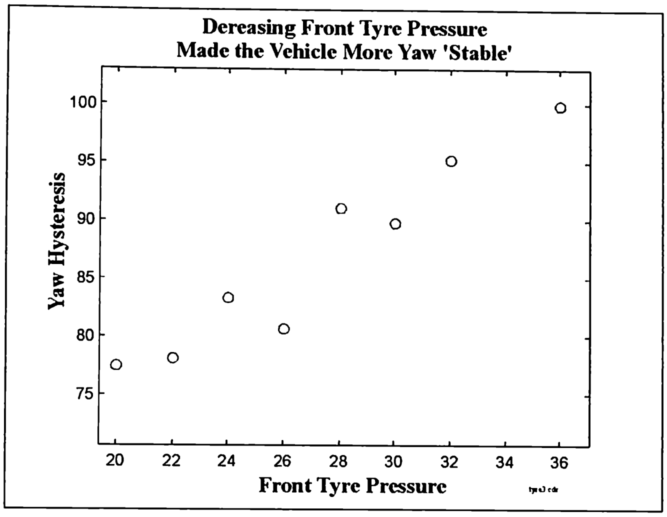


Figure 94

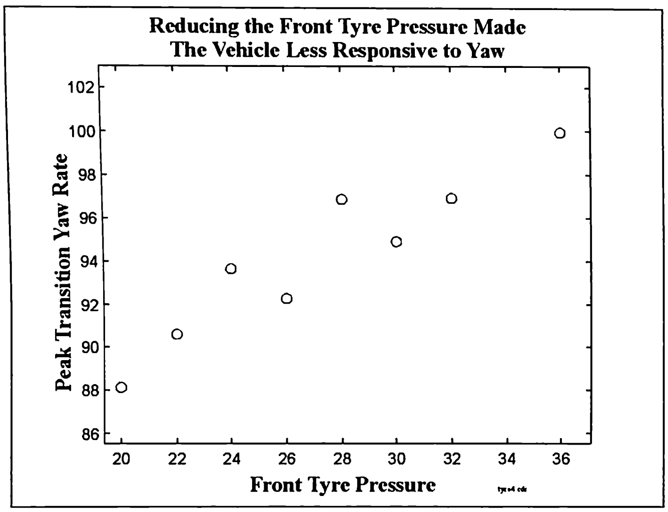


Figure 95

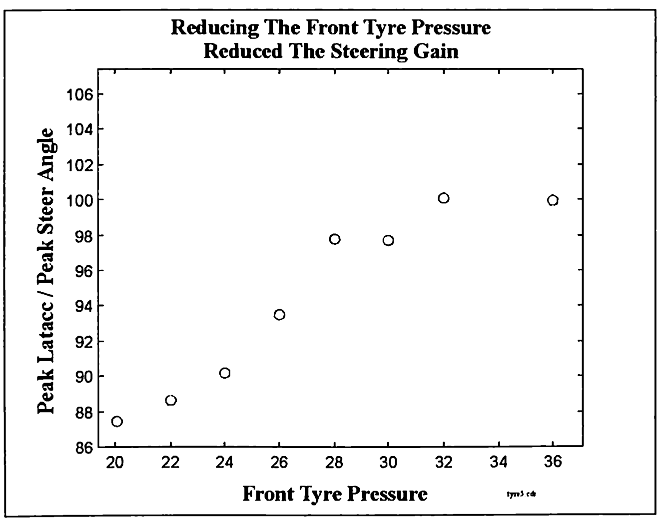


Figure 96

In conclusion, altering the front tyre pressures produced a systematic and well behaved change in the vehicle's behaviour that was of a large enough amplitude that it should have been detected by a good VDE. Furthermore, the change was robust across different driving styles and altered the secondary behaviours of the vehicle in a coherent fashion.

8.3.2 Effect of Training on Discrimination Thresholds

Figure 97 shows the number of subjects who performed at each level of discrimination performance before and after training. Prior to training most subjects were in the worst discrimination performance category (<=20psi) whereas after training most subjects were placed in the best category (>=32psi).

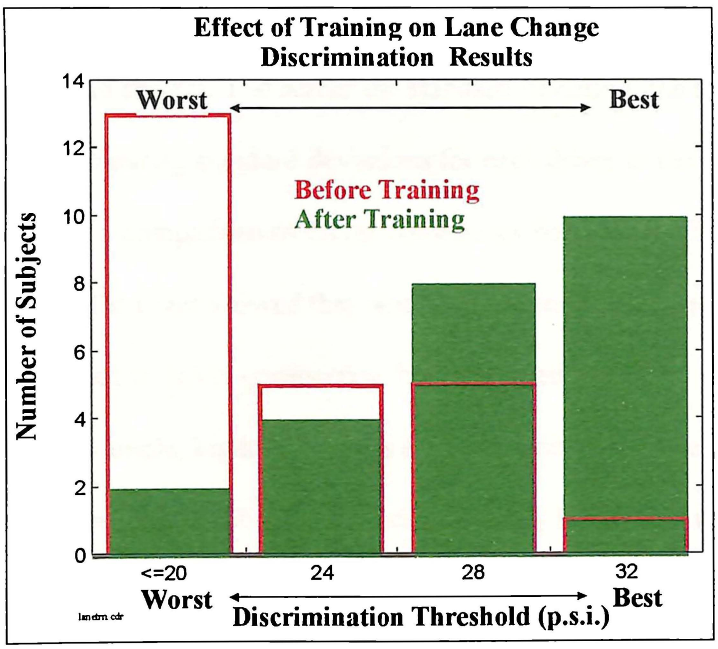


Figure 97

Figure 97 shows that both the distributions (before and after training) were not normally distributed about their means but were highly skewed (Pearson first coefficient of skewness = +0.79 before training and -0.94 after training). The explanation for this skewness is that tyre pressures lower than 20 psi were not tested because Ford management were concerned about tyre / rim separation at lower pressures which could lead to an accident. Therefore the distribution is truncated on the left for the pre-training distribution. Tyre pressures above 34 psi could not be tested without using smaller increments and hence the post training distribution is truncated to the right. This extreme skewness makes the application of the t-test of significance less reliable but that test indicates that there was a significant difference between the untrained and trained discrimination results ($p < 0.001$). Thus training produced a highly significant improvement in discrimination performance.

8.3.3 Effect of Driver Training on Driver Consistency

Driver consistency was measured by determining the standard deviation of each of the 36 metrics for all drivers. The higher the standard deviation, the more inconsistent the driver was. Comparing standard deviations for each driver in the two discrimination sessions allows a comparison of driver consistency before and after training. Application of the t-test showed that when the drivers were treated as a group there was no difference in driver consistency before and after training for every metric ($p > 0.1$). For example, Figure 98 shows a comparison of driver consistency before and after training for each of the 24 drivers calculated by finding the mean of the standard deviations of the 5 most highly correlated metrics (from section 8.3a above). Figure 98 shows that some drivers were less consistent after training while other drivers were more consistent after training.

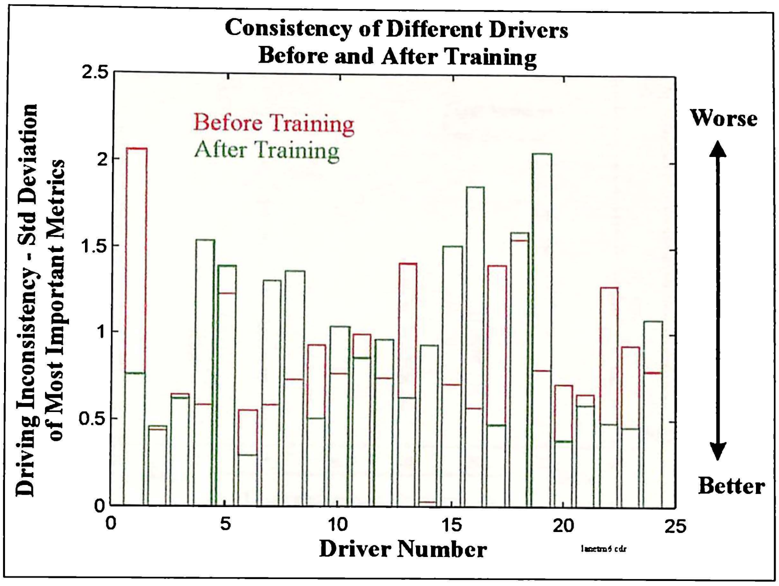


Figure 98

8.3.4 Relation of Driver Consistency to Discrimination Performance

A correlation analysis determined that there was no correlation between driver consistency and discrimination performance in either the first discrimination session, the second discrimination session or for both sessions combined for all 36 metrics ($p>0.1$). That is, consistent drivers were no better at discriminating changes to the vehicle than inconsistent drivers. For example Figure 99 shows the relation of the standard deviation of Yaw Rate hysteresis (a measure of driving consistency) against discrimination performance for each of the 24 drivers, before and after they received training. The correlation coefficient for this data is established and plotted as a single value on Figure 100, which displays the results obtained for all other driver consistency metrics and discrimination performance.

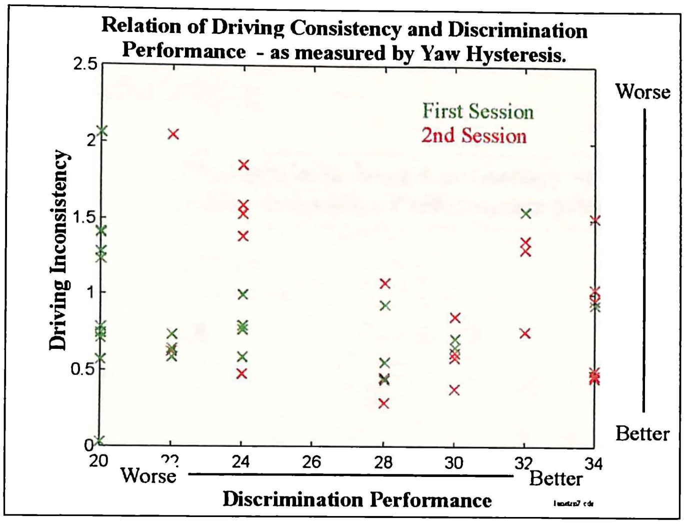


Figure 99

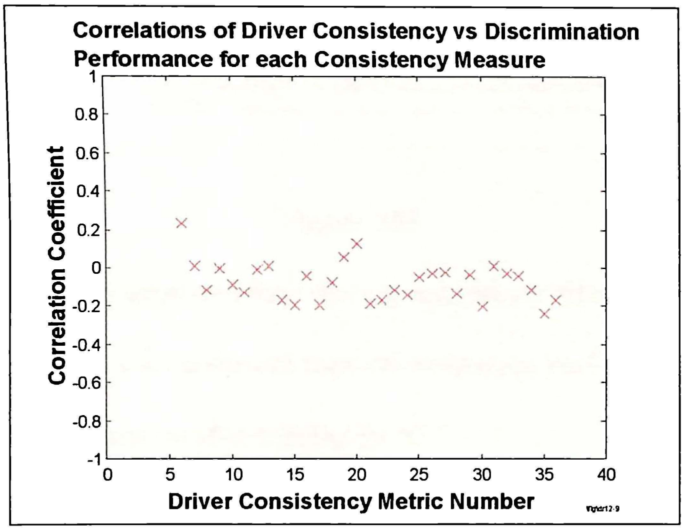


Figure 100

8.3.5 Changes in Driver Consistency and Changes in Discrimination Performance

Further analysis found that changes in discrimination performance after training were not related to changes in driving consistency ($p>0.1$). For example, Figure 101 below shows the changes in discrimination performance on the x-axis for all 24 drivers after training while the y-axis shows whether the subject improved or decreased their driving consistency as measured by the change in their standard deviation of the metric Yaw Rate hysteresis after training.

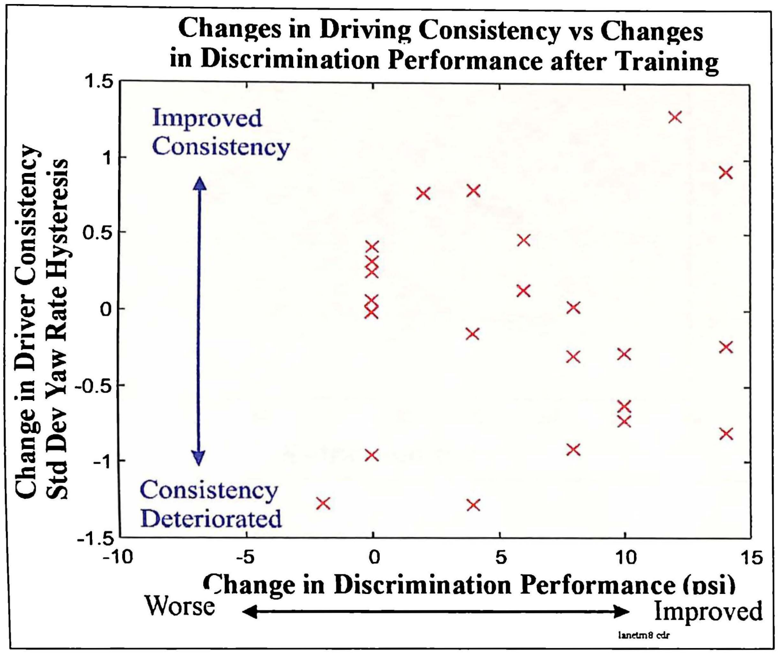


Figure 101

Thus VDEs who improved their driving consistency after training were no more likely to be the drivers who improved their discrimination performance than VDEs who became less consistent after training ($p > 0.1$).

8.3.6 Relation of Driving Style to Discrimination Performance

Figures 92 and 93 above showed that there were large differences in driving style. A series of correlation analyses showed that there was no relation between driving style and discrimination performance ($p > 0.1$) either before training, after training or for both sessions combined. Figure 102 shows the correlation coefficients relating mean driving style to discrimination performance for each of the driving style metrics.

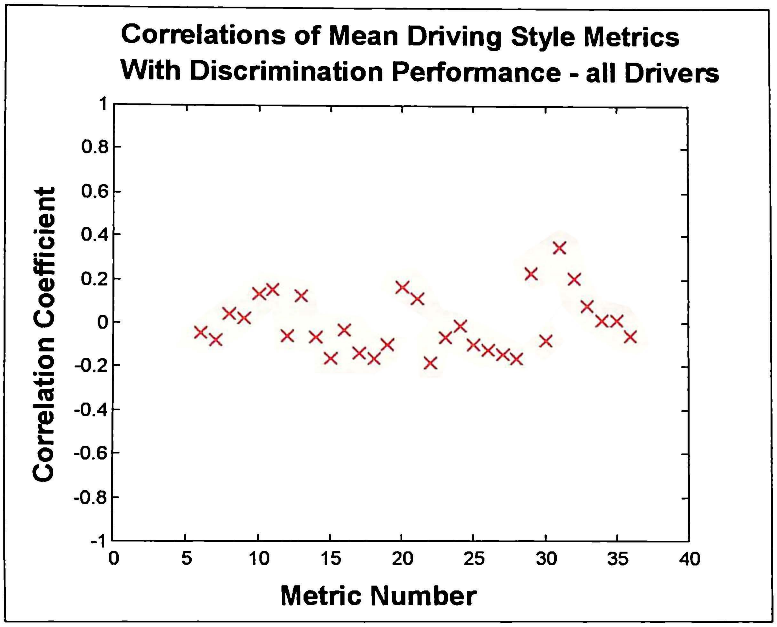


Figure 102

8.3.7 Subjective Rating of Training Course

After each VDE had completed both days of the Lane Change discrimination course they were asked to rate the effectiveness of the course in helping them improve their evaluation skills. Using the 0 - 10 scale listed below, the average rating was 9.9 with a mode of 10.

0 = Very poor, of no value whatsoever

2 = Poor, of not much value

4 = Reasonable value

6 = Above average course - of good value

8 = Very good course - of great value

10 = Exceptional Course.

8.4 Discussion

The overall poor performance of the VDE's prior to training but after self-practice was again both a surprise and a concern to senior Ford management. However both the VDEs themselves and their senior managers were pleased with the improvements that resulted after training. Prior to this investigation, the prevailing belief at Ford was that self practice alone in sufficient quantities would suffice to allow each subject to reach their discrimination performance limit. As many of the subjects performed the Lane Change manoeuvre as part of their daily job, it was therefore believed that many of them were already at that limit. The expectation was that training interventions would yield little or no improvement in performance particularly as the number of Lane Change trials conducted during this investigation for each subject was a small fraction of the total number of Lane Change trials that they had performed during their career. Rather, management at Ford expected good VDEs to possess 'Attributes' or 'Innate Skills' that were correlated with high levels of discrimination performance that distinguished them from lesser performers. It was hoped that identification of these attributes would allow a program of **selecting** better VDE candidates to be implemented. Instead our investigation showed the exact reverse. Some of the worst performing subjects with many years of prior experience who consistently failed at the easiest test could be trained within a few hours to perform reliably at a level equal to that of Jackie Stewart. Conversely, it was also found that there was no correlation between the RAF Attributes tests nor the vestibular tests with Lane Change discrimination performance as we shall see later in section 10 below. This leads to the conclusion that vehicle evaluation performance (at least at this level of expertise) as measured by the Lane Change discrimination test is primarily a learnt skill rather than an innate ability. Furthermore, unaided practice appears to be a poor method of learning this skill as most subjects demonstrated poor performance after a large quantity of self practice over many years but demonstrated good performance after a small amount of training.

The finding which was most surprising to both the VDEs and ourselves was that driver consistency was not related to discrimination performance. After all, it was the lack of driving consistency demonstrated at Florida in 1993 that initiated this entire program of VDE investigation. If neither driving consistency nor driving style was related to discrimination performance, then what was? This question is of vital concern to any ongoing training program because we need to understand the factors that limit discrimination performance in order to be able to develop the most effective training program. Section 9 investigates this question in detail.

9 Factors Affecting Lane Change Discrimination Performance

9.1 Introduction

Traditional methods of Signal Detection Theory (Peterson, Birdsall and Fox, 1954; Green and Swets, 1974; van Meter and Middleton, 1954) allow us to determine **if** there has been an increase in stimulus discriminability, as distinct from the subject's criteria for reporting discrimination, but usually it does not allow us to investigate **why** such an increase might have occurred. For example, Figure 103 below shows two hypothetical Receiver Operating Curves (ROC) for a discrimination task.

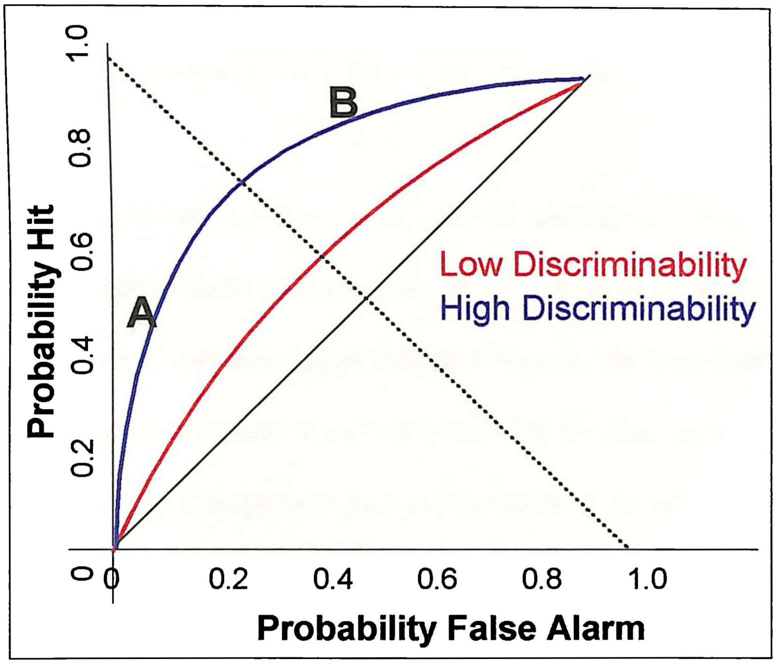


Figure 103 Signal Detection Theory Receiver Operating Curves

In Figure 103 the blue ROC curve is generated for 2 stimuli which are more discriminable than those generated by the red ROC curve. Normally the area between the ROC curve and the solid diagonal line is defined as the measure of the stimulus discriminability and is referred to as d' although sometimes other measures such as maximal displacement from the solid diagonal line are used (Green and Swets, 1974). Different points along each curve correspond to different subjective decision strategies. For example the subject has a less strict criteria at point A on the blue curve than at point B while the discriminability of the stimulus remains the same.

However, if a 2 Alternate Forced choice procedure is used as in our Lane Change Discrimination test and there is no temporal response bias, then the ROC curve will simply be defined by a single point which lies on the dashed diagonal line with the abscissa being P1,ns and the ordinate being P1,sn (Falmagne,1990), where;

i) P1,sn = the probability that the subject reports the first stimulus was larger

ii) P1,ns = the probability that the subject reports the 2nd stimulus was larger

If any temporal response bias does occur, this will generate two points symmetrically distributed around the dashed diagonal line and still result in the same ROC curve. (Falmagne, 1990). Combining Signal Detection Theory, the 2 alternate forced choice methodology and the Staircase algorithm means that the discrimination thresholds we measured in the Lane Change were an actual measure of the subjects' stimulus discriminability which was independent of their decision criteria. Section 8 above showed that after training these discrimination scores improved. This section investigates the underlying factors that gave rise to this improvement.

An analysis of the tasks performed by a VDE during the Lane Change discrimination test shows that they broadly fall into 2 categories;

i) Control of Vehicle Motion - e.g. Driving style or Driving Consistency

ii) Perception of Vehicle Motion. - eg sensory transduction, encoding, recall etc.

A subject's improved discrimination ability could be due to improvements in either or both of these categories. For example, by controlling the vehicle more consistently it could be that one subject has a more reliable estimate of the Base stimulus (36 psi) with which to compare the Modified stimulus (say 24 psi) than another driver who is so inconsistent that their repeated experience of the Base is never the same, or it could be that one driving style amplifies the differences between the base and the modified runs allowing better discrimination. But sections 8.3d, 8.3c and 8.3f also showed us that driver Control consistency and style were not correlated with discrimination performance. This raises the possibility that the increase in discrimination performance was due to an increase in Perceptual performance. Two questions then arise;

i) How do we determine if Perceptual performance was actually improved by training

ii) Were Perceptual improvements responsible for Discrimination improvements

We cannot simply infer improvements in Perceptual performance from our failure to establish a correlation between Control style or Control consistency and discrimination performance as it could well be that there are other measures of Control performance that we have neglected to measure or indeed that there is another entire category of VDE performance separate from the two that we have postulated above. To answer these questions we need a method that separates out the influence of Control Performance from the influence of Perceptual Performance on Discrimination performance. This requirement is addressed below by the development of a mathematical model which results in a family of Stimulus Discriminability vs Driver Confidence curves which can be transformed to a family of 3-dimensional ROC curves that generate a sensitivity surface in space.

This mathematical model produces analytical techniques that can be used in a training program to help identify perceptual difficulties that a VDE may be experiencing. In fact these techniques were developed and used during the current investigation.

9.2 Mathematical Model of Driver Control and Perceptual Performance

A 4 parameter mathematical model of driver Control performance and Perceptual Performance was constructed to simulate some of the processes involved when a driver executes the Lane Change. These parameters were varied during a Monte Carlo simulation and the results compared with actual driver data. The 4 parameters are;

i) The VDE's Control Consistency

ii) The VDEs Perceptual Noise due to Transduction, Encoding, Recall, Comparison

iii) The VDE's Subjective Sensitivity Rating Scale

iv) The Vehicle's set up condition - tyre pressure

In the above 4 parameter model, item i) is thus a measure of Control performance, items ii) and iii) are measures of Perceptual performance and item iv) describes the experimental condition. The reason that we have not included a 5th term to describe the driver's Control Style is that our analysis in section 8.3a (Figures 91, 92 and 93) showed that the differences between the Base and the Modified behaviour of the vehicle is to a large extent independent of the driving style. That is the stimulus d' for violent or gentle manoeuvres is statistically identical for equal changes in tyre pressure. The terms i), ii) and iii) are defined in detail later.

The schematic outline of each iteration of the Monte Carlo simulation is shown below in Figure 104. In this simulation, for each setting of the 3 parameters i) - ii) above that we have chosen, we model 1,600 pairs of Lane Change trials consisting of 200 Base and 200 Modified runs at each of the Modified tyre pressures from 20 to 34 psi (in 2 psi increments). Each time a new combination of parameters is chosen we repeat the 1600 trials.

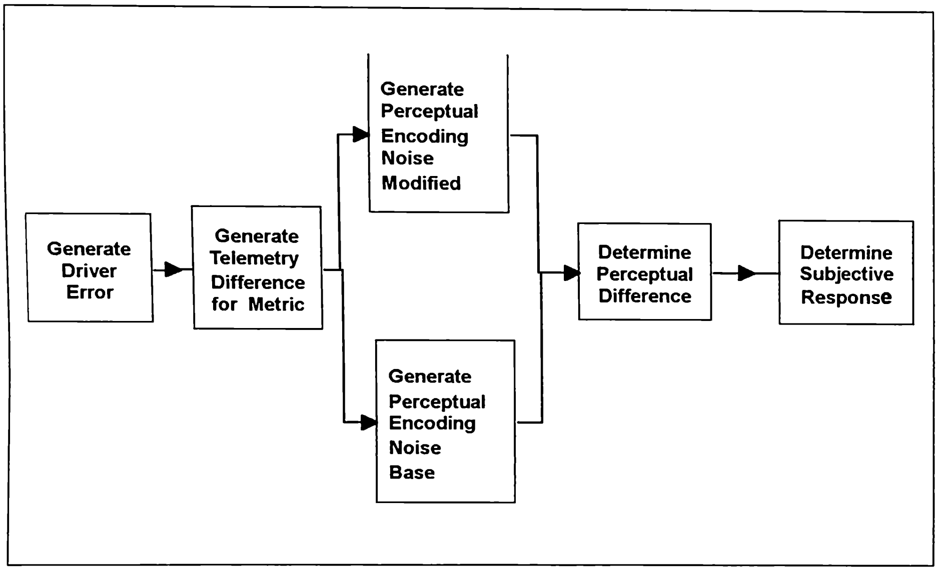


Figure 104 Flow Chart of a Single Iteration of the Monte Carlo Simulation

9.2.1 Modelling Control Consistency

In Section 8 we saw that changes to the tyre pressure settings produced well behaved changes to the objective behaviour of the vehicle that the subjects were using to base their discrimination on. In particular, Figures 91, 94 - 96 showed that there was a well behaved mapping of tyre pressure settings to the mean value of each objective metric. This means that for any driver, we can convert the tyre pressure setting to an equivalent objective measure of vehicle behaviour. Thus a tyre pressure setting of 26 psi might correspond to a mean value of Transition Steer integral of 550 while a tyre pressure setting of 28 psi might correspond to a mean Transition Steer integral of 510. Different drivers will have a different mapping because of their difference in control style, but the relative difference between these objective measures or the discriminability of these differences will be the same [2*(550 - 510)/(550+510)]. Because this mapping is a monotonic function, we can <u>reverse the process</u> and determine for any run what the <u>equivalent mean tyre pressure</u> would be for that run irrespective of the actual tyre pressures used. For example, if we measured the Transition Steer integral on one run as 510 then we would say that the vehicle behaved as it normally would with tyre pressures set to 28 psi. That is, the steering responsiveness of the vehicle in this run was equivalent to an average run at 28 psi., even if the vehicle had tyre pressures set at say 24 psi. The reason for the discrepancy between the actual tyre pressures and the equivalent tyre pressures is that the driver may have driven the vehicle in an unusual fashion which made the 24 psi setting behave as if it was a 28 psi setting. For example, we have seen in Figures 48 and 49 that increasing the entry speed or the peak entry Steer Angle will produce an increase in transition steer angle from that normally expected. Thus any driving style can be converted to a <u>psi equivalent</u> and hence any driving inconsistency can be converted to an <u>equivalent standard deviation in psi settings</u>. Therefore the Driving Consistency parameter i) can be specified in psi equivalents.

Figure 105 shows how we simulated the drivers' Control consistency.

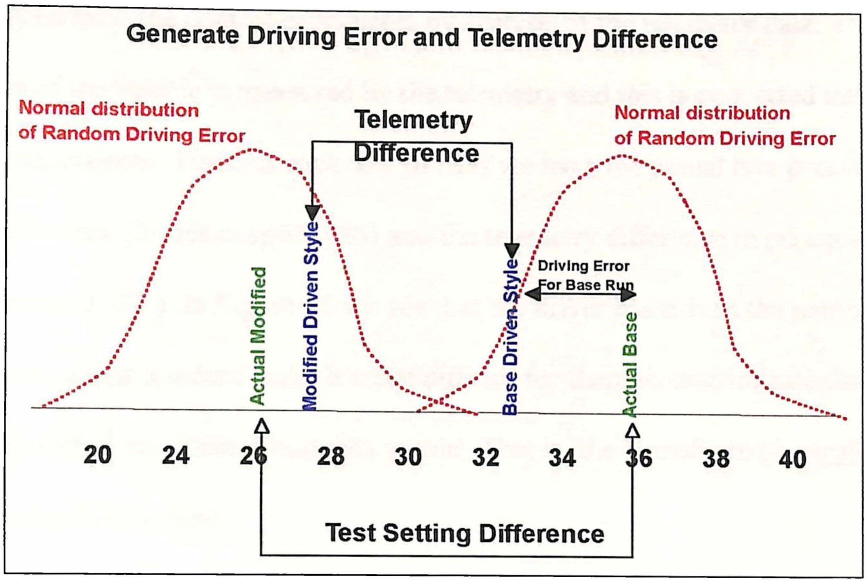


Figure 105

In Figure 105 the driver has been given a pair of trials to discriminate with the Base tyre pressures actually set at 36 psi and the Modified tyre pressures set at 26 psi. Because the driver makes random errors in their driving, sometimes entering too fast, sometimes using different amounts of steering input at the entry section, the objective measure of the vehicle's behaviour when transformed into psi equivalents, will be normally distributed about the actual psi settings. This is indicated by the red distributions about each actual tyre pressure setting. The mean of the distributions is set at the actual tyre pressure used for the run while the probability distribution shows the likelihood that the driver has driven the vehicle in such a fashion that it behaves as if the vehicle had the pressures indicated on the x-axis. The larger the standard deviation of the red distributions, the more inconsistent the driver's control is.

In this particular pair of runs the driver has driven the Base run so that it behaved as it would normally do at 33 psi and the Modified run as it normally would do if set to 27 psi. We determine the control equivalents by analysis of the telemetry data. The actual behaviour of the vehicle is measured by the telemetry and this is converted into pressure equivalents. Thus for each pair of runs we have the actual tyre pressure setting difference (in this case 36 - 26) and the telemetry difference in psi equivalents (in this case 33 -27). In Figure 98 we see that the driver has driven the pair of runs in such a fashion that it would make it more difficult for them to discriminate the Base and the Modified runs than it normally would. That is, the stimuli are physically less discriminable than normal.

In summary, driver control consistency can be characterised by a standard deviation in psi equivalents and can be measured objectively by the telemetry.

9.2.2 Modelling Perceptual Noise

Now that we have produced the objective behaviour of the vehicle, we have to model how the subject perceives that behaviour. Figure 106 below shows how the Perceptual process was modelled.

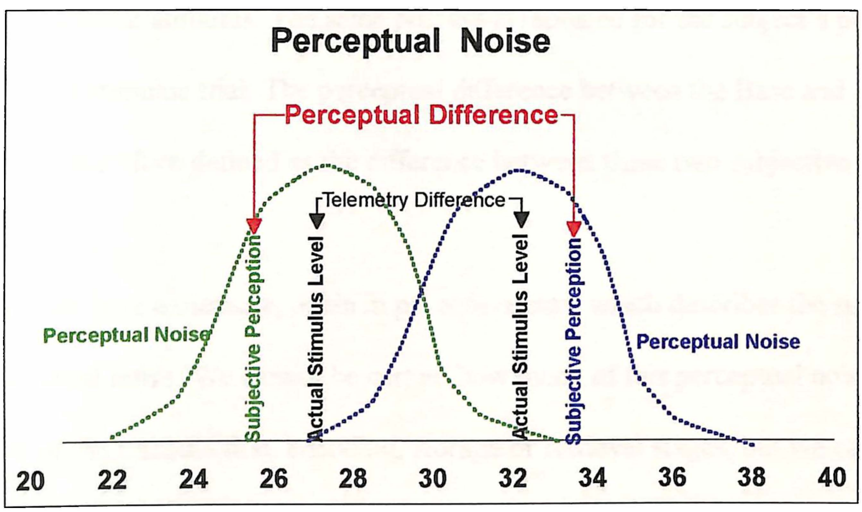


Figure 106 Modelling the Perceptual Encoding Noise

The actual behaviour of the vehicle as determined by the telemetry difference from Figure 105 is carried over to Figure 106. This is because the physical stimuli that the subject must discriminate are those measured by the telemetry and not determined by the vehicle settings. It is generally accepted that a given stimulus level repeatedly given to a subject produces an effect on the subject's sensory apparatus that is randomly distributed about a mean value. This model is referred to the Random Utility Model (Block and Marschak, 1960; Luce and Suppes, 1965; Marschak, 1960; Falmagne, 1978; Mc Fadden and Richter, 1970, 1971; and Manski 1977) and forms the underlying basis of most psychophysical theories. This random distribution is shown in Figure 106 by the two gaussian curves centred about the telemetry values. Again we can use the concept of psi equivalents to mathematically describe this random perceptual variation. The larger the standard deviations of these perceptual distributions, the more noisy we say that the subject's perceptual processes were. In the example above the subject was presented with a Modified stimulus that behaved as if it was from a normal vehicle setting of 27 psi, but the subject perceived that stimulus due to these random variations in their perceptual processes as if it was normally a 25 psi equivalent stimulus. The same process is repeated for the subject's perception of the Base stimulus trial. The perceptual difference between the Base and the Modified trials is therefore defined as the difference between these two subjective perceptions.

Thus we have a measure, again in psi equivalents, which describes the subject's perceptual noise. We cannot be certain how much of this perceptual noise arises at each of the transduction, encoding, storage or retrieval stages, but we can describe the overall noise of the entire system without determining these component contributions.

9.2.3 Modelling the Subjective Sensitivity Rating Scale

Now that we have determined the perceptual difference that the subject experiences between the two runs we must convert this into a Confidence Rating. This is achieved by the use of a scaling factor as shown below in Figure 107. The larger the scaling factor the higher the subject's confidence for any given level of Perceptual difference.

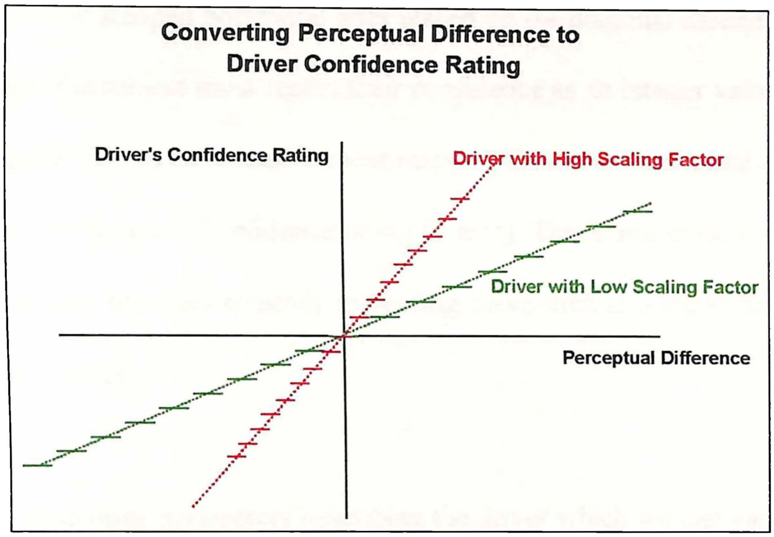


Figure 107

In Figure 107 the 'red driver' has a higher scaling factor than the 'green driver' and thus for any given level of Perceptual difference presented to these two drivers, the red driver will rate with a higher confidence than the green driver. This scaling difference between the red and the green drivers could arise because of two possibilities or combinations of them;

i) The red driver is actually more sensitive to the sensory stimuli than the green driver. ie., small changes in the sensory stimulus feel stronger to the red driver than the green driver

ii) The red driver is less conservative than the green driver in their rating scale i.e., the red driver might actually 'feel' the same difference between the two stimuli but uses a different interpretation as to what a confidence 8 for example means to them

Notice in Figure 107 that the subject's responses are quantized into integers which is depicted by the stepped horizontal lines placed on the diagonal dashed lines as in our experiment, the subject must report their confidence as an integer value between 0 and 10. In Figure 107 we have used a linear mapping from the Perceptual difference (x-axis) to the Driver's Confidence rating (y-axis). This is not critical and the model can be run with any monotonically increasing curve such as a logarithm function, a power function etc.

Thus we have three parameters describing the driver which we can vary;

i) Driving Consistency

ii) Perceptual noise

iii) Confidence Scaling factor

The question that we wish to address is whether we can develop a method for separating out the effect of these three parameters in our Lane Change Discrimination test. This was achieved by conducting the Monte Carlo simulations.

9.3 The Stimulus Discriminabilty vs Driver Confidence Graph

The Monte Carlo simulations will be evaluated using a graph called the 'Stimulus Discriminability vs Driver Confidence' graph or SDDC graph for short. Each telemetry metric generates its own unique SDDC graph which is constructed in the following manner. The x-axis plots the driver's Confidence that they are correct in their discrimination evaluation. If the driver is incorrect in their evaluation then their Confidence value is multiplied by -1 before being plotted. For each Confidence value from - 10 to +10, we find all the Base and all the Modified runs corresponding to that rating and calculate the mean telemetry values for the metric in question for those Base and Modified runs. The difference between these means is plotted for each Confidence value on the y-axis. A typical SDDC graph for a subject is plotted below in Figure 108.

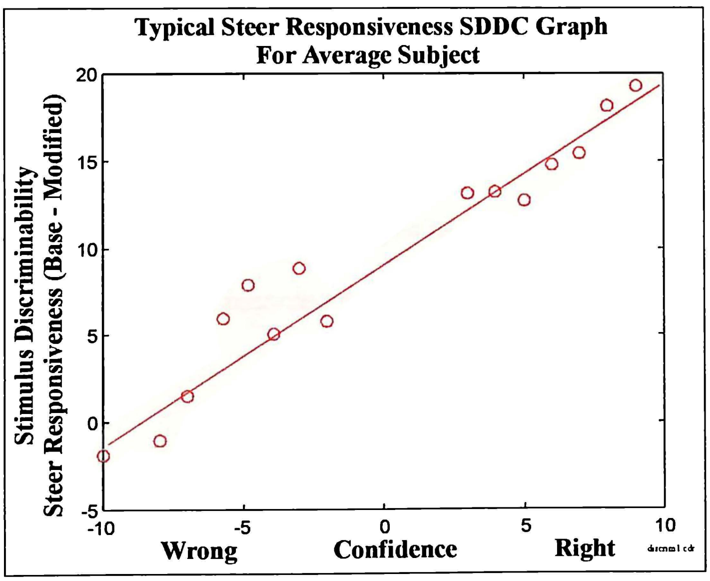


Figure 108

Figure 108 shows that as the difference between the Base and the Modified Steer Responsivenes Metric increases, the subject is more likely to be correct with high confidence. As this difference becomes smaller, the subject responds with lower confidence and then starts being on average wrong as indicated by the negative confidence ratings. At the extreme bottom left of Figure 108 we see that the behaviour of the vehicle in the Base conditions and the Modified conditions has reversed. That is the mean telemetry difference which is normally positive is now negative. In this case the driver has driven the Base vehicle in such a fashion that it behaved more like the Modified vehicle and vice versa.

The important point to notice about Figure 108 is that the y-axis is a composite of the vehicle setting (tyre pressure) and the driver's control style and consistency. For any difference between the Base and the Modified runs as measured by the telemetry, we do not know what the vehicle setting was nor how the driver executed that run. We only know what was the physical sensation (as measured by the telemetry) that the driver experienced.

We will generate a number of these SDDC graphs in our Monte Carlo simulations and compare them with actual graphs recorded for the VDEs during the Lane Change discrimination test.

9.4 The Monte Carlo Simulations

The Monte Carlo simulations were run by holding 2 of the parameters a) - c) constant while varying the remaining one and the experimental parameter d) to determine the effect of each parameter on the SDDC graph.

9.4.1 Effect of Driver Control Consistency on SDDC graph

The effect of driver Control Consistency was determined by arbitrarily setting;

i) The Sensitivity Scaling Factor to a constant = 0.75

ii) The Perceptual Noise to a constant = +/- 1.0 psi equivalents.

While the driver Control Consistency was varied in 4 sets of simulation runs;

a) std deviation of +/- 0.1 psi equivalents
b) std deviation of +/- 1 psi equivalents
c) std deviation of +/- 2 psi equivalents
d) std deviation of +/- 4 psi equivalents

Each simulation run consisted of 400 trials. This produced the SDDC graph shown in Figure 109 below.

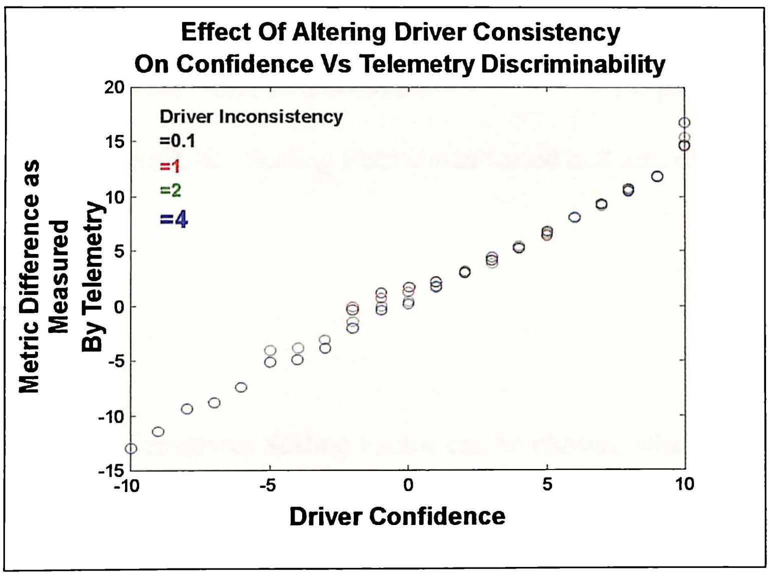


Figure 109

Figure 109 shows that large changes in the driver's Control Consistency produced no significant effect on the shape of the curve. In fact the curves can be made arbitrarily close by increasing the number of trials in the Monte Carlo simulation. This result is entirely expected and could be derived from first principles if need be. The reason for this is that altering the driver's control consistency simply alters the physical stimulus that the subject will experience. That is, it simply chooses different values from the y-axis and does not influence the relation between the axes as explained in section 9.2 a) - c) above. This was one of the motivations behind deriving the SDDC graph in the first place. We wanted to define a measure which was independent of both the driver's Control consistency and their Discrimination Performance. These trials were performed simply to provide an audit of the Monte Carlo software.

9.4.2 Effect of Sensitivity Scaling Factor on the SDDC graph

The effect of altering the Sensitivity Scaling Factor on the SDDC graphs was determined by arbitrarily setting;

i) The Control Consistency to a constant = + / - 1.0 psi equivalents.

ii) The Perceptual Noise to a constant = + / - 1.0 psi equivalents.

While the driver Sensitivity Scaling Factor was varied in 4 sets of simulation runs;

a) 0.5
b) 0.75
c) 1.25
d) 2.0

Any values of the Sensitivity Scaling Factor can be chosen, what we are interested in is the qualitative effect on the shape of the SDDC graph.

Each simulation run consisted of 400 trials. This produced the SDDC graph shown in Figure 110 below.

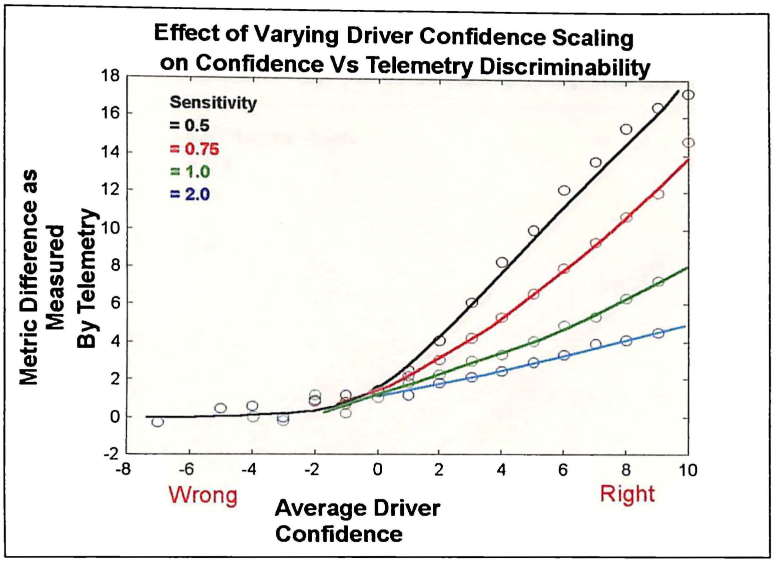


Figure 110

Figure 110 shows that decreasing the Sensitivity Scaling Factor primarily produces an increase in the gradient of the SDDC graph to the right of the Confidence = 0 line.

9.4.3 Effect of Varying Perceptual Noise on the SDDC graph

The effect of varying Perceptual Noise on the SDDC graph was determined by arbitrarily setting;

i) The Sensitivity Scaling Factor to a constant = 0.75

ii) The Control Consistency to a constant = + / - 1.0 psi equivalents.

While the Perceptual Noise was varied in 4 sets of simulation runs;

a) std deviation of +/- 0.1 psi equivalents
b) std deviation of +/- 1 psi equivalents
c) std deviation of +/- 2 psi equivalents
d) std deviation of +/- 4 psi equivalents

Each simulation run consisted of 400 trials. This produced the SDDC graph shown in Figure 111 below.

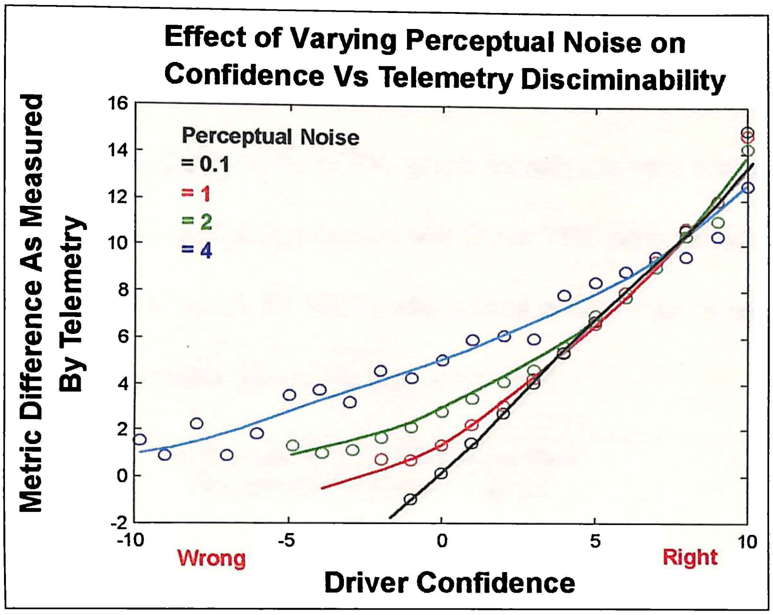


Figure 111

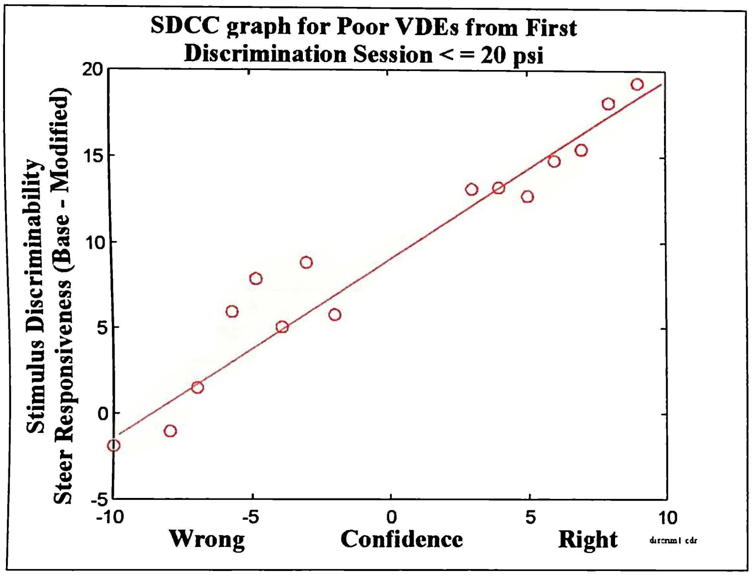
The primary effect of increasing the Perceptual Noise is an increase in the vertical offset of the graph above and to the left of the confidence = 0 line. With low Perceptual Noise the graph almost passes through the origin (0,0) whereas with high Perceptual noise the x intercept is at y ~= 4. In fact it can be shown that at x= 0, y is equal to the Perceptual Noise.

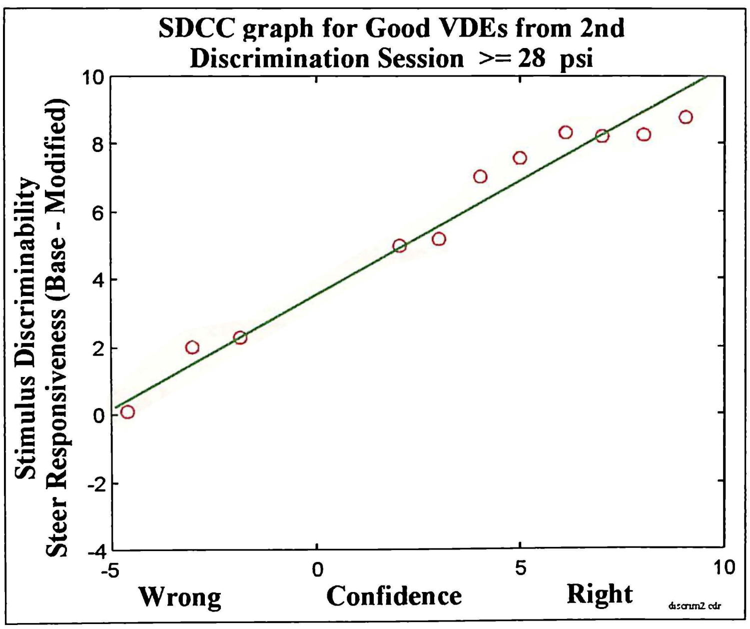
Each of the features in graphs 109 - 111 can be made significant to any level required by extending the sample size of the Monte Carlo simulation. With our sample size of 1600 trials per simulation, all differences are significant at $p < 0.001$.

We are now in a position to evaluate the SDDC graphs arising from the Lane Change Discrimination Test.

9.5 SDDC graphs from Lane Change Discrimination Testing

To determine what the effect of training was on the subjects, two SDDC graphs were generated. Figure 112 shows the SDDC graph for subjects who scored less than or equal to 20 psi in the first discrimination test (Poor VDE performance) while Figure 113 shows the SDDC graph for VDEs who scored greater than or equal to 28 for the second discrimination test (Good VDE performance).





Figures 112 & 113

In both Figures 112 and 113 the linear best fit lines were fitted using the method of least squares (Spiegel, 1961). For comparison purposes these best fit lines are plotted together in Figure 114 below.

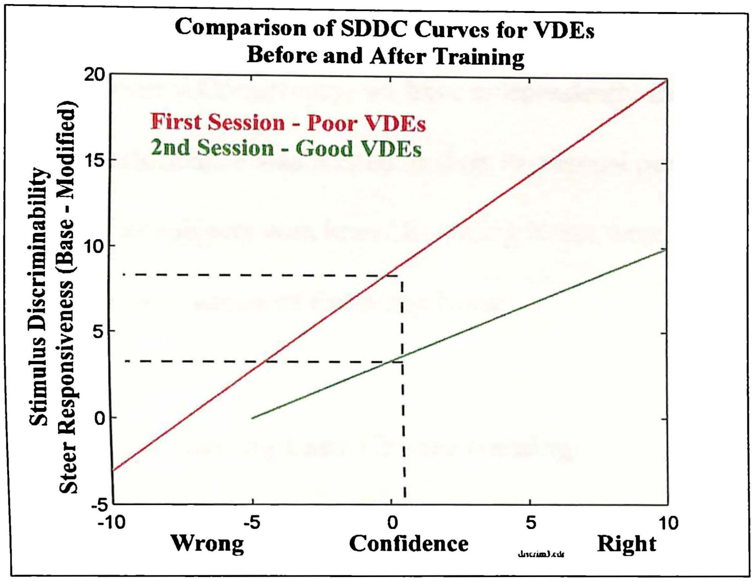


Figure 114

Figure 114 shows that the main differences between the two curves are;

i) The increased y value for at x=0 for the Poor VDEs (y= 3 vs y = 9)

ii) The increased gradient for the Poor VDEs. (grad = 0.7 vs 1.1)

From section 9.4c above we know that the change in y values at x=0 corresponds to a change in the subject's Perceptual noise. In particular, our linear regressions suggest that the good VDEs have approximately 3 times less Perceptual noise than the Poor VDEs ($p<0.001$) and hence should have markedly superior Perceptual performance near threshold. Further, the reduction in gradient for the good VDEs corresponds to being more conservative in their Confidence rating for a given sensory experience.

That is, the good VDEs required a greater sensory difference before they would give a high confidence rating than the Poor VDEs.

Thus not only have we shown that the VDEs' Discrimination performance was not related to their Control Consistency, we have independently shown that their Discrimination performance was related to their Perceptual performance. In particular we have shown that subjects with lower Encoding Noise were better evaluators than subjects who had large values of Encoding Noise.

9.6 Use of SDDC graphs during Lane Change training

The SDDC graphs and their underlying mathematical analysis were not just theoretical developments. They proved to be most valuable tools for evaluating and training the VDEs. This section demonstrates how they were used. In a number of cases during the first Discrimination test, the subject would score randomly (ie. approximately 50% right and 50% wrong) but would do so with repeatedly **high** confidence ratings. The subject clearly believed that they felt something markedly different between the runs, but whatever it was they felt, was obviously not correlated with the tyre pressure changes. Otherwise they would either have scored correctly or systematically incorrectly by reliably reversing their identification of the Base and the Modified runs. In this case the subject may have been concentrating on feeling something which was not directly related to the tyre pressures. Once this pattern of responding became apparent, the Discrimination test was terminated and a training program commenced. The first thing that had to be established in such a training session was, 'What sensation, if anything, were the subjects actually responding to?'

In many cases at this early stage of the course, the subjects had little actual idea of exactly what it was that they were basing their responses on. They just tended to lump all their perceptions into a single overall impression. The SDDC graph was originally developed to help answer this question. If the subject was basing their responses on a particular aspect of the vehicle's behaviour, then there should be a correlation between the metric summarising this vehicle behaviour and their Confidence scores, independent of the subject's driving style and the tyre pressure settings. The computer was programmed to identify which of the vehicle behaviours the subject was responding to and which they were ignoring, by generating the SDDC graphs for each metric and establishing the correlation coefficients for them. The higher the correlation coefficient, the more likely it was that the subject was basing their evaluation on that metric. For example Figure 115 below shows a metric which is not correlated with the driver's decision criteria and hence would be rejected as one which was upsetting the driver. Notice the contrast between this graph and the previous SDDC graphs.

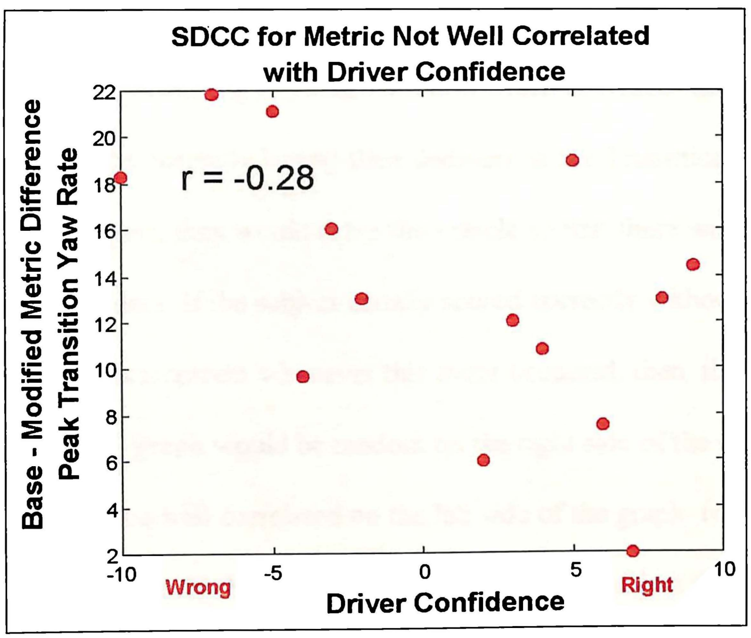


Figure 115

Once the metric which was correlated with the driver's responding was identified, training could then proceed constructively by showing the subject that this metric was not correlated to the tyre pressure settings. This would be followed by practical driving demonstrations where the subject was shown how to identify the correct sensations during the manoeuvre. In this way the driver's model of the vehicle's dynamic behaviour was developed by helping them to understand which metrics were most affected by tyre pressure changes and which metrics were most affected by different driving actions. The relation of vehicle set-up and /or driving style to vehicle behaviour defines the subject's model of the vehicle's dynamics. Furthermore, this process refined the subject's focus of perceptual attention during the execution of the Lane Change.

The same graph was also useful for helping a subject who in general was scoring accurately and conservatively, but who had the occasional error with high confidence. It would often turn out that these occasional high confidence errors were caused by some infrequently occurring physical sensation which confused the driver. For example the driver might be correctly basing their decision on the Transition Steer Integral, but every now and again, they would drive the vehicle so that there was a sharp Yaw Rate overshoot in the data. If the subject usually scored correctly without the Yaw Rate overshoot but was incorrect whenever this event occurred, then the Yaw Rate Overshoot SDDC graph would be random on the right side of the graph (confident and right), but would be well correlated on the left side of the graph (confident and wrong). Training would then include identifying for the subject these infrequent sensations that were upsetting their accuracy and helping them to *'tune out'* these sensations as being unreliable or confusing. Thus the subject would be made aware of the effect of the large Yaw Rate overshoot on their accuracy and be instructed to say to themselves, 'Ignore this Yaw Rate sensation, it often confuses me. Instead I should just remember what the Transition Steer Integral felt like, even although it was of a smaller magnitude, it is more reliable.' This procedure worked remarkably well for curing such a problem with these subjects and extended their knowledge of the vehicle's dynamic behaviour.

9.7 Conclusion

A mathematical model which simulated driver Perceptual performance and Driver Control performance was constructed. This model explained the key features observed in SDDC graphs derived from actual VDE evaluation sessions. Application of the analytical techniques arising from this model to the Ford VDEs' data showed that improvements in their evaluation performance were primarily derived from improvements in their Perceptual performance rather than in their Control performance. This result ties in nicely with the results obtained during the vestibular experiments. In that series of investigations it was found that subjects could improve their discrimination performance by making use of techniques such as Focus Restriction or Special Encoding algorithms.

The SDDC graph proved to be a powerful tool during training and allowed the instructor to identify the subjects' weaknesses. It is expected that these analytical tools could be applied to other disciplines where skilled performance is dependent on combinations of both perception and control.

10 R.A.F. Basic Attributes Test

10.1 Introduction

The RAF Basic Attributes tests attempt to measure innate cognitive attributes which are independent of the subject's training history and which predict the subject's final trained potential at performing various tasks that are deemed to be important for success in operating fighter aircraft. (Caretta,1989; Turnbull, 1992; Caretta, 1992). Long term studies by the Ministry of Defence (UK) and the DRA have shown that their Basic Attributes tests have the following predictive values for pilot success in the following different aircraft industries;

Test Type	C.V.T	S.M.A.	Combined Weighted Scores
R.A.F.	0.33	0.33	0.42
Turkish A.F.	0.25	0.2	0.32
Army Air Corps	0.2	0.42	0.45
Qantas	0.24	0.32	n/a

Table 6 Correlation Coefficients of RAF Basic Attributes Scores with Flight Graduation Success

(Directorate of Recruitment and Selection (RAF) - Internal Document PTC/152100/R&S. W/6Aptitude.)

Discussion with the RAF indicates that the primary purpose of these tests is to 'eliminate duffers' rather than to 'select the elite'. That is, the tests are primarily used to help identify those candidates that are likely to fail the training program, rather than to identify those candidates that are likely to perform exceptionally well in the program.

As stated above, one of the original aims of the VDE investigation conducted at Upper Heyford was to determine which if any Attributes were correlated with vehicle evaluation performance (Lane Change Discrimination performance). Therefore each subject's BAT scores were correlated with their discrimination performance.

10.2 Method

10.2.1 Procedure

All subjects were transported to the Directorate of Recruitment at Cranwell and tested by the Cranwell staff as if they were prospective Air Force fighter pilots. The exact methodology used by the Directorate of Recruiting is confidential to them and we are not permitted to publish this. Nevertheless we can report that all tests were conducted using a special computerised interface which contained a keyboard, joystick, foot pedals, and special response buttons. The test was administered automatically and scored by the computer.

10.2.2 Tests Used

The tests administered were;

i) Control of Velocity Test (C.V.T)

Hand - Eye co-ordination

This test implements an Anticipatory tracking task. A pattern of red dots scroll down the screen at a constant rate. The candidate controls a white dot positioned three quarters of the way down the screen. The dot can be moved horizontally in both directions to intercept the red dots by using the joystick. The more red dots intercepted the higher the score.

ii) **Digit Recall Test**

Short Term Memory

A simple memory test. The test displays a number on the screen for a short period of time. The candidate must remember the number and then type it in after a delay. The length of the numbers increases to find the maximum size the subject can remember.

iii) **Ravens Matrices Mat- 62**

General Intelligence

Mat-62 is a widely used measure of non-verbal intelligence. The test is made up of 36 questions with 1 question used as a worked example. Each question consists of a 3 x3 matrix of symbols with the bottom right symbol missing. The candidate must choose the correct missing symbol from 8 alternatives.

iv) **Scheduling Test**

Multi-tasking

The test presents many screens each showing a bar growing horizontally across the screen against a scale of 1 to 100. The candidate must stop the bar as close to 100 as possible by pressing a designated key. Each screen is only visible by pressing its number on the keypad. At the start of the test only one screen is active, but this increases to nine screens by the end of the test. The candidate must 'schedule' their viewing of the screens in order to stop each individual bar at the optimum time.

v) **Sensory Motor Apparatus**

Complex Hand - Eye - Foot Co-ordination

This test implements a Compensatory tracking task. A single spot moves around the screen following a programmed path. The candidate must try to keep the spot over the middle of a cross-hair at the centre of the screen. The candidate controls the spot using a joystick and foot pedals. The fewer the excursions from the centre of the screen the higher the score.

vi) **Vigilance**

Speed / Accuracy

This test comprises a matrix in which asterisks (routine task) and arrows (priority task) appear. Asterisks are cancelled from the screen by inputting the row and the column. If an arrow appears, this must be cancelled immediately by initiating a priority task and then cancelling using the same technique as asterisks. Three measures are taken from this test and are reported here and in the next two test results. For this test result speed of response to the priority task is measured.

vii **Vigilance 1**

The accuracy of the cancellation (row / column co-ordinates) for test vi) above.

viii) Vigilance 2

The number of asterisks cleared for test vi) above.

ix) Spatial Reasoning

In this test a cube with markings on each of its side is shown with 4 of its faces flattened out in plane view. The subject is presented with 6 other cubes similarly represented but in difference orientations and with the markings moved around on the faces. The subject has to choose which of the six cubes could be rotated to match the initial cube.

These tests can be summarised as measuring the following attributes;

1) Control Precision

a) Hands Only

b) Multi- Limb

2) Reasoning

a) Patterns and Relationships

b) Visualisation and Orientation

3) Information Processing

a) Short term memory capacity

b) Monitoring Changes

c) Controlling simultaneous tasks

10.3 Results

The RAF made available to us;

i) The Raw scores for each test

ii) The scores for each test converted to a stanine (1-9 category with 9 being the best)

iii) The overall weighted scores for all tests combined. The scores were weighted so that they gave the highest statistical correlation with historical pilot success

Because some raw scores are better when smaller and others are better when larger, the most meaningful presentation of the results is by using their stanine categories. Figure 116 shows the comparison of the overall stanine ratings for the 24 VDEs compared to the RAF pilots.

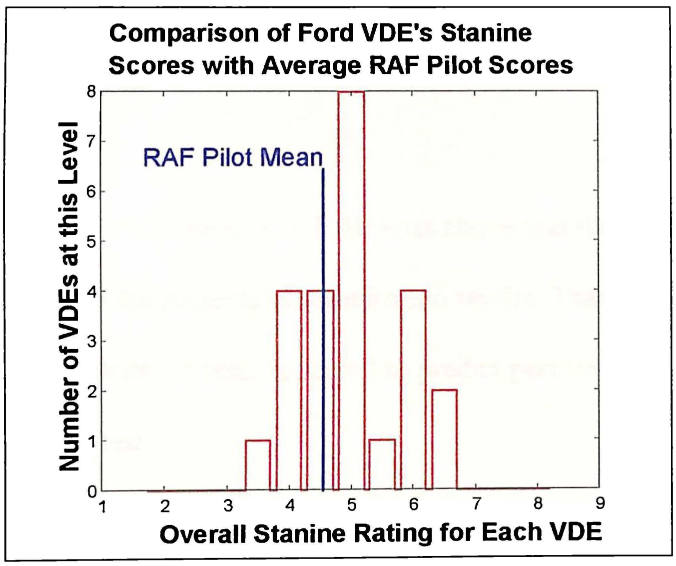


Figure 116

Because no standard deviations are available for the RAF pilots, a test of significance cannot be performed. However Figure 116 tends to indicate that the VDEs were at least on a par with the average pilot applicant.

A comprehensive statistical analysis was conducted to see if the RAF Attributes tests correlated with any aspect of Lane Change Discrimination Performance.

The

i) Raw scores for each individual test

ii) Stanine scores for each test

iii) Overall weighted scores for all tests combined for each subject

iv) Scores for tests combined by Attribute type

were correlated with each subject's;

i) First Discrimination Test Result

ii) Second Discrimination Test Result

iii) Best Discrimination Test Result

iv) Average Test Result

It was found that for all combinations of RAF tests above that there was no significant correlation ($p>0.1$) with the subjects' discrimination results. That is no test or combination of RAF Attributes tests appeared to predict performance in the Lane Change Discrimination test.

10.4 Discussion

It is important to remember that the population sample that the VDE's are being compared to is already highly selected and no way representative of the general population. The group norm for subjects sitting the Basic Attributes tests are young males aged 17-25, highly motivated, with good academic qualifications who have already passed stringent RAF pre-selection procedures. Thus a subject who scores 4 on the stanine rating for any test, while being only average with respect to the pilot applicants, is well above the general non-fighter pilot population. Thus the VDEs scored well in the Attributes test.

The failure to establish any correlation with the VDEs Lane Change discrimination performance could have arisen because of at least 2 reasons;

i) The subjects did not reach their upper limit of discrimination performance during the Lane Change test but if they had then the BAT scores would have been correlated

ii) The BAT scores do not measure the attributes that are critical for predicting Lane Change discrimination performance at the required level

Unfortunately we are not in a position to determine which of these or other hypotheses are likely to be true. However the fact that such strong improvements in discrimination performance were achieved by training and the fact that no significant correlations were established with the BAT scores leads us to favour conclusion b) for the time being.

In a large scale meta analysis of 68 published studies, comprising 468 correlations from a sample of 437,258 subjects during 1940 to 1990, Hunter and Burke (1994) found that when all the results were combined the most consistently correlated variable with obtained validity was decade of publishing with a decline in average validities over the 5 decades studied. That is, the main conclusion was that later studies showed less validity for the Attributes tests than the earlier studies.

Hunter and Burke rated the Attributes measures according to a 4 category class of validity based on Hunter and Schmidt's (1990) Validity Generalizability and Whitener's 90% credibility limit. The four categories were;

i) 75% rule positive and 90% confidence limit positive. No true variation in studies. Mean Validity generalisable.

ii) 75% rule negative and 90% confidence limit positive. Validity non zero but size of validity influenced by moderators.

iii) 75% rule negative and 90% confidence limit negative. Uncertainty about true validity of predictor. Validity not generalisable.

iv) 75% rule positive and 90% confidence negative. True validity is zero.

It was found that no tests satisfied category i), that is were generalisable. Spatial ability ($r = 0.109$), perceptual speed ($r=0.29$) and reaction time ($r = 0.18$) satisfied category ii) but their correlation coefficients were low. The main reason promoted by Hunter and Burke as to why decade of publication correlated with a decline in validity was that in later studies the subjects were more highly pre-selected. With a more highly selected group it has been found that correlations are reduced as the greatest contribution to the predictive power of the tests is caused by the range exhibited between the best and the worst subjects. For example, Thorndike's (1949) validities for 7 tests ranged from 0.18 to 0.46 for unselected candidates and combined to give a composite validity of 0.64. Later a cut-off criteria was used for selecting candidates (top 13%) and when this cut-off criteria was used retrospectively on the original sample, it was found that the composite validity fell to 0.18. In particular the psychomotor validity fell from 0.46 in the unselected group to -0.03 in the selected group. Figure 116 shows that if modern RAF fighter pilot candidates are highly pre-selected, then the VDEs are at least as pre-selected.

Thus it may be that with such highly selected subjects in our VDE pool that Attributes tests may be of little value as increasingly appears to be the case with airforce pilots. That is, the tests primarily eliminate' duffers' but do not select the 'elite' from the good.

10.5 Conclusion

The major emphasis for the future should be on developing training techniques to improve existing VDEs rather than on trying to establish selection criteria.

11 Overall Conclusion of VDE Investigation

It was found that unaided practice and traditional Ford training procedures did not produce VDE performance of the standard required by the modern motor industry. On the contrary, the VDEs were highly confident in their ability even when they repeatedly failed on tests that were directly related to their daily job requirements. As a group the VDEs were already highly selected as far as Attribute skills were concerned but this selection was not sufficient to produce good VDE performance.

A new mathematical model was developed which described the VDEs' Control and Perceptual performances. From this model, analytical techniques were developed which allowed the instructor to identify weaknesses in a subject's Perceptual or Control performance and evaluate the effectiveness of specific training interventions. These analytical techniques may turn out to have applications in other areas of human performance research.

A training program was conducted with the VDEs and it was found that the primary factor limiting the VDEs' vehicle evaluation performance was their Perceptual performance rather than their Control performance, contrary to initial expectation.

It was found that the VDEs could be trained by altering the techniques they used to perceive and encode motion and that these techniques produced substantial improvements in both their vestibular discrimination test results and their Lane Change discrimination test results. The successfulness of the Focus Restriction technique used to enhance the VDEs' Perceptual performance was explained in terms of a Vestibular Short Term Sensory Store which briefly contained more information than the subjects were typically able to process when attending to the entire stimulus presentation.

These investigations have therefore ushered in a new era in Motor Vehicle development with Ford Motor Co. committing to a 5 year program based on these psychological training techniques.

In a separate study, similar results were found with elite race drivers. That study is outlined in section 12 below and detailed fully in the Attachment. Traditional methods of training and telemetry analysis only produced improvements for subjects up to the elite level. At the elite level these techniques provided no further improvement. A new type of analysis was developed which produced significant improvements even for elite drivers. This has culminated in a future program (1997/1998) with a Formula 1 team and the largest supplier of telemetry equipment to develop the next generation of mathematical analysis tools.

12 Summary of Elite Race Driver Investigations

A completely independent set of investigations was undertaken with elite race drivers. While these investigations were independent and focused on different skills, it became apparent that many of the techniques that were useful in the VDE investigations were also useful in the Elite Race Driver investigations. This section contains an overview of the findings that arose from the Elite Race driver investigations. The full study is contained in Attachment A.

i) A training program was conducted in Florida during May 1992 with Jackie Stewart who was used as a benchmark to train his son Paul. At that time Paul was competing in the European F-3000 championship, the level immediately below Formula 1. Despite being retired for more than 20 years and being in his mid 50's, Jackie was always the fastest driver. During this training session, it was found that there was a highly visible fault in Paul's control of the vehicle when compared to Jackie's control. This fault was corrected using telemetry analysis and feedback which resulted in Paul's lap times then being approximately equal to Jackie's.

ii) In the process of correcting Paul's large control fault, we also subtly altered the shape of his curvature through the corners. It was not clear how much of Paul's improvement in performance was due to curing his control fault or whether the alteration of the curvature profiles had some impact on his lap times. Further, it was found that Jackie always operated the car slightly further away from its performance limit than Paul did and yet Jackie still managed to achieve the fast laps. This was surprising and so I developed a series of computer simulations during 1993 and 1994 which modelled the effects of altering curvature profiles on lap times (Spackman, K.P. & Tan, M.S, 1993).

iii) It was found that the curvature profile used by Jackie closely matched the optimal solution produced by the computer simulations. The computer simulations also showed that different vehicles would require different solutions for the same corner. Conversely, different corner and straight combinations required different solutions for the same vehicle.

iv) When tested in different vehicles on the same track it was found that Jackie also adjusted his curvature profiles to match the vehicle in the same manner that the computer simulations did. This alteration of curvature profiles to match the track and vehicle combination seemed to be something unique to Jackie. Other highly skilled racing drivers seemed to find a particular curvature solution and then adhere to this solution even when the vehicle was changed. It was as if these drivers did not have an algorithm that allowed them to optimise their curvature control. This finding was sufficiently novel that it was published in New Scientist.

v) Unfortunately, the computer simulations for any particular corner and vehicle combination are so time consuming and complex, that it is not practical to calculate the optimal solution in real racing situations. Slight changes in track temperature, vehicle set-up etc. require weeks of new computer modelling. However, the computer models did provide the theoretical understanding of how altering the curvature interacted with other factors affecting lap times. I therefore decided to approach the problem from the reverse direction by using empirical techniques. By defining metrics which summarised how the driver actually controlled the vehicle during a race, it was possible to perform statistical analyses of these metrics with lap times to determine what were the optimal curvature profiles under the current experimental conditions. This technique was termed Metric Meta Analysis.

vi) A further study was conducted at Pukekohe in January 1995 with;

a) The then current and 2 times World Touring Car champion, Paul Radisich.

b) Five times Australian Touring Car champion, Peter Brock.

c) Three times American IMSA GTP champion, Steve Millen.

d) A 19 year Junior NZ Formula Ford Champion, Jason Richards.

These four drivers knew the Pukekohe track intimately as they had just completed a 1 week international meeting there. Further, the 3 champions had previously competed at the track more than 70 times between them. Each driver drove the same instrumented vehicle until they could no longer improve on their lap times.

vii) Analysis of the telemetry data found that all 4 drivers had arrived at completely different curvature solutions for the same corner. Metric Meta Analysis showed that none of these solutions was the optimal one.

viii) The current World Champion and the 5 times Australian Champion were significantly faster than the American champion who was in turn consistently faster than the N.Z. Junior Champion, as expected. It was decided to conduct a program with the N.Z. Junior Champion (Jason), using telemetry feedback to train him to drive the vehicle in the manner which the Metric Meta Analysis predicted would be the optimal solution. After 3 practice and feedback sessions totalling 3 hours in a single afternoon, Jason was able to operate the vehicle in the fashion prescribed by the computer.

This new driving style produced 2 important results;

i) Jason produced lap times which were 1.3 seconds faster than any of the 3 International Champions had ever managed to do.

ii) Jason subjectively felt that this style of driving was much slower than his natural style and he 'Couldn't believe I had asked him to drive in this fashion'.

ix) The conclusion from the above studies was that most champion drivers do not possess the correct information processing algorithm which would allow them to determine the optimal curvature control solution from the incoming sensory information. The drivers appear to be tricked by a form of 'motion illusion' in much the same way that people are deceived by visual illusions. Their interpretation of the world does not correlate correctly with the actual physical world.

x) By utilisation of correct training procedures it is possible to get the subjects to not only execute the optimal solution, but to also alter their mental model so that it now feels faster to them as well.

References

Aczel, J. (1966)

Lectures on Functional equations and their Applications. New York: Academic Press.

Alpern, M., Rushton, W. A. H., & Tori, S. (1970a)

The attenuation of rod signals by backgrounds. *Journal of Physiology,* 206, 209-227.

Alpern, M., Rushton, W. A. H., & Tori, S. (1970b)

The signals from cones. *Journal of Physiology,* 207, 463-475.

Ambler, R. K., Bair, J. T., & Wherry, R. J. (1960)

Factorial structure and validity of naval aviator selector variables. *Aerospace Medicine,* 31, 456-461.

Ambler, R. K., Johnson, C. W., & Clark, B. (1952)

An analysis of biographical inventory and Spatial Apperception Test scores in relation to other selection tests (Special Report No. 52-5). Pensacola, FL: U.S. Naval School of Aviation Medicine.

Anderson, N. H. (1970a)

Averaging model applied to the size-weight illusion. *Perception and Psychophysics,* 8, 1-4.

Anderson, N. H. (1970b)

Functional measurement and psychophysical judgement. *Psychological Review,* 77, 153-170.

Anderson, N. H. (1974)

Information integration theory: A brief survey. In Krantz, D. H., Atkinson, R. C., Luce, R. D., & Suppes, P., (Eds.), *Contemporary Developments in Mathematical*

Psychology, Vol. 2: Measurement, Psychophysics, and Neural information processing. San Francisco: Freeman, W. H.

Anderson, N. H. (1976)
Integration theory, functional measurement and the psychophysical law. In Geissler, H. G., & Zabrodin, Y. U. M., (Eds.), *Advances in Psychophysics.* Berlin: VEB Deutscher Verlag.

Anderson, N. H. (1981)
Foundations of Information Integration Theory. New York: Academic Press.

Anderson, T. W., McCarthy, P. I., & Tukey, I. W. (1946)
Staircase methods of sensitivity testing. *Navord Reports,* 46-65.

Arth, T. O., Steuck, K. W., Sorrentino, C. T., & Burke, E. F. (1988)
Air Force Officer Qualifying Test (AFOQT): Predictors of undergraduate pilot training and undergraduate navigator training success (Tech. Rep. No. AFHRL-TP-88-27). Brooks Air Force Base, TX: Air Force Human Resources Laboratory.

Bair, J. T., Lockman, R. F., & Martoccia, C. T. (1956)
Validility and factor analysis of naval air training predictor and criterion measures. *Journal of Applied Psychology,* 40, 213-219.

Baird, J. C. (1970)
Psychophysical Analysis of Visual Space. Oxford: Pergamon Press.

Baird, J. C., Green, D. M., & Luce, R. D. (1980)
Variability and sequential effects in cross-modality matching of area and loudness. *Journal of Experimental Psychology: Human Perception and Performance,* 6, 277-289.

Baker, R. A., & Osgood, S. W. (1954)

Discrimination transfer along a pitch continuum. *J. exp. Psychol.*, 48, 241-246.

Bale, R. M., & Ambler, R. K. (1971)

Application of college and flight background questionnaires as supplementary noncognitive measures for use in the selection of student naval aviators. *Aerospace Medicine*, 42, 1178-1181.

Bartram, D., & Dale, H. C. A. (1982)

The Eysenck Personality Inventory as a selection test for military pilots. *Journal of Occupational Psychology*, 55, 287-296.

Bebie, H., Frankhauser, F., & Spahr, J. (1976)

Static perimetry: Strategies. *Acta Ophthalmologica*, 54, 325-338.

Benson, A. J. (1990)

Sensory function and limitations of the vestibular system. *In Perception and Control of Self Motion*, 143-210. Edited Warren, R. and Wertheim, A. published Lawrence Erlbaum Associates.

Benson, A. J., & Brown, S. F. (1989)

Visual display lowers detection threshold of angular, but not linear, whole body motion stimuli. *Aviat. Space and Env. Med.*, 60, 629-633.

Benson, A. J., Hutt, E. C. B., & Brown, S. F. (1989)

Thresholds for the perception of whole body angular movement about a vertical axis. *Aviat. Space and Env Med*, 60, 205-213.

Benson, A. J., Kass, J. R., & Vogel, H. (1986)

European vestibular experiments on the spacelab-1 mission:4: Thresholds of perception of whole-body linear oscillation. *Exp. Brain Res.*, 64, 264-271.

Benson, A. J., Spencer, B. A., & Stott, J. R. R. (1986)

Threshold for the detection of the direction of whole body linear movement in the horizontal plane. *Aviat., Space and Env. Med.*, 57, 1088-1096.

Berkshire, J. R. (1967)

Evaluation of several experimental aviation selection tests. Pensacola, FL: Naval Aerospace Medical Center.

Berkshre, J. R., & Ambler, R. K. (1963)

The value of indoctrination flights in the screening and training of Naval aviators. *Aerospace Medicine,* 34, 420-423.

Berliner, J. E., & Durlach, N. I. (1973)

Intensity perception IV. Resolution in roving-level discrimination. *Journal of the Acoustical Society of America,* 53, 1270-1287.

Berliner, J. E., Durlach, N. I., & Braida, L. D. (1977)

Intensity perception IV. Futher data on roving-level discrimination and the resolution and bias edge effects. *Journal of the Acoustical Society of America,* 61, 1577-1585.

Birnbaum, M. H. (1978)

Differences and ratios in psychological measurement. In Restle, F., & Castellan, jr, N. J., (Eds.), *Cognitive Theory,* 3, Hillsdale, N. J.: Erlbaum.

Birnbaum, M. H. (1982)

Controversies in psychological measurement. In Wegener, B., (Eds.), *Social attitudes and Psychophysical Measurement.* Hillsdale, N. H.: Erlbaum.

Bjorkman, M., & Ottander, C. (1959)

Improvements of discriminative ability by training. *Reports from the Psychol. Lab.*, University of Stockholm, No66.

Benson, A. J., Spencer, B. A., & Stott, J. R. R. (1986)
Threshold for the detection of the direction of whole body linear movement in the horizontal plane. *Aviat., Space and Env. Med.*, 57, 1088-1096.

Berkshire, J. R. (1967)
Evaluation of several experimental aviation selection tests. Pensacola, FL: Naval Aerospace Medical Center.

Berkshre, J. R., & Ambler, R. K. (1963)
The value of indoctrination flights in the screening and training of Naval aviators. *Aerospace Medicine,* 34, 420-423.

Berliner, J. E., & Durlach, N. I. (1973)
Intensity perception IV. Resolution in roving-level discrimination. *Journal of the Acoustical Society of America,* 53, 1270-1287.

Berliner, J. E., Durlach, N. I., & Braida, L. D. (1977)
Intensity perception IV. Futher data on roving-level discrimination and the resolution and bias edge effects. *Journal of the Acoustical Society of America,* 61, 1577-1585.

Birnbaum, M. H. (1978)
Differences and ratios in psychological measurement. In Restle, F., & Castellan, jr, N. J., (Eds.), *Cognitive Theory,* 3, Hillsdale, N. J.: Erlbaum.

Birnbaum, M. H. (1982)
Controversies in psychological measurement. In Wegener, B., (Eds.), *Social attitudes and Psychophysical Measurement.* Hillsdale, N. H.: Erlbaum.

Bjorkman, M., & Ottander, C. (1959)
Improvements of discriminative ability by training. *Reports from the Psychol. Lab.*, University of Stockholm, No66.

Blake, R., & Fox, R. (1973)

The psychophysical inquiry into binocular summation. *Perception and Psychophysics,* 14, 161-185.

Block, H. D., & Marschak, J. (1960)

Random orderings and stochastic theories of responses. In. Olkin, Ghurye, Madow, Mann and Hoeffding (Eds.), *Contributions to Probability and Statistics.* Stanford University press.

Block, R. D., & Jones, L. V. (1968)

The Measurement and Prediction of Judgement and Choice. San Francisco: Holden-Day.

Blum, J. R. (1954)

Approximation methods which converge with probability one. *Annals of Mathematical Statistics,* 25, 382-386.

Bordelon, V. P., & Kantor, J. E. (1986)

Ulitization of psychomotor screening for USAF pilot candidates: Independent and integrated selection methodologies (Tech. Rep. No. AFHRL-TR-86-4). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Boring, E. G., Langfield, H. S., & Weld, H. P. (1948)

Foundations of Psychology. New York Wiley Press.

Bradley, R. A. (1954a)

Incomplete block rank analysis: On the appropriateness of the model for a method of paired comparisons. *Biometrics,* 10, 375-390.

Bradley, R. A. (1954b)

Rank analysis of incomplete block designs. II. Additional tables for the method of paired comparisons. *Biometrika,* 41, 502-537.

Bradley, R. A. (1955)

Rank analysis of incomplete block designs. III. Additional tables for the method of paired comparisons. *Biometrika,* 42, 450-470.

Bradley, R. A., & Terry, M. E. (1952)

Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika,* 39, 324-345.

Braida, L. D., & Durlach, N. I. (1972)

Intensity perception. II. Resolution in one-interval paradigms. *Journal of the Acoustical Society of America,* 51, 483-502.

Bramber, D. (1975)

The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *Journal of Mathematical Psychology,* 12, 387-415.

Broadbent, D. E., & Gregory, M. (1968)

Vigilance considered as a statistical decision. *British Journal of Psychology,* 54, 309-323.

Brown, J., & Cane, V. R. (1959)

An analysis of the limiting method. *British Journal of Statistical Psychology,* 12, 119-126.

Brugger, P., Landis, T., & Regard, M. (1990)

A "sheep-goat effect" in repetition avoidance: Extra sensory perception as an effect of subjective probability? *British Journal of Psychology,* 81, 455-468.

Burke, E. F. (1980)

Results of a preliminary study on a new tracking test for pilot selection (Note No. 9/80). London: Ministry of Defence, Science 3 (Royal Air Force).

Carretta, T. R. (1987)

The Basic Attributes Tests: An experimental selection and classification instrument for U.S. Air Force pilot candidates. In R.S. Jensen (Ed.), *Proceedings of the Fourth International Symposium on Aviation Psychology* (pp. 500-507). Columbus: Ohio State University, Aviation Psychology Laboratory.

Carretta, T. R. (1989)

USAF pilot selection and classification systems.. *Aviat. Space Environ. Med.*, 60, 46-49.

Carretta, T. R. (1992)

Recent developments in US air force pilot candidate selection and classification. *Aviat. Space Environ. Med.*, 63, 1112-4.

Carretta, T. R., & Siem, F. M. (1988)

Personality, attitudes, and pilot training performance: Final analysis (Tech. Rep. No. AFHRL-TP-88-23). Brooks Air Force Base, TX: Air Force Human Resources Laboratory , Manpower and Personnel Division.

Carroll, J. B., & Arabie, P. (1980)

Multidimensional scaling. *Annual Review of Psychology*, 31, 607-649.

Cherry, E. C. (1953)

Some experiments on the recognition of speech, with one and two ears. *Journal of Accoustical Society of America*, 25, 975-979.

Chocolle, R. (1940)

Variations des temps de reaction auditifs en fonction de l'intensite a diverses frequences. *Anne Psychologique*, 41, 65-124.

Clark, B. (1970)

The vestibular system. *Annual Review of Psychology*, 21, 273-306.

Cliff, N. (1973)

Scaling. *Annual Review of Psychology*, 24, 473-506.

Cohen, J. (1977)

Statistical power analysis for the behavioral sciences. New York: Academic.

Cohen, J. (1983)

The cost of dichotomization. *Applied Psychological Measurement*, 7, 249-253.

Collett, D. (1991)

Modelling Binary Data. London: Chapman & Hall.

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970)

A Mathematical Psychology: An Elementary Introduction. Englewood Cliffs, N. J.: Prentice-Hall.

Cornsweet, T. N. (1962)

The staircase method in psychophysics, *Am. J. Psychol.*, 75, 485-491.

Cornsweet, T. N., & Pinsker, H. M. (1965)

Luminance discrimination of brief flashes under various conditions of adaption. *Journal of Physicology*, 17 (6), 713-719.

Cox, R. H. (1988)

Utilization of psychomotor screening for USAF pilot candidates: Enhancing predictive validity. *Aviation, Space & Environmental Medicine,* 59, 640-645.

Cramer, H. (1963)

Mathematical Methods of Statistics. Princeton, N. J.: Princeton Unversity Press, 1963.

Creager, J. G. (1983)

Human Anatomy and Physiology. Belmont. Wadsworth.

Croll, P. R., Mullins, C. J., & Weeks, J. L. (1973)

Validation of the cross-cultural aircrew aptitude battery on a Vietnamese pilot trainee sample (Tech. Rep. No. AFHRL-TR-73-30). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, Personnel Research Division.

Cross, D. V. (1965)

An application of mean value theory to psychological measurement. In *Progress Report No.* 6 (Report No. 05613-3-P). Ann Arbor. The behavioral analysis laboratory, University of Michigan.

Daintith, J., & Nelson, R. D. (1989)

Dictionary of Mathematics. London: Penguin Books.

Damos, D. L., & Lintern, G. (1979)

A comparision of single - and dual-task measures to predict pilot performance (Engineering Psychology Tech. Rep. No. 79/2). Urbana-Champaign: University of Illinois at Urbana-Champaign.

Dantzig, G. B. (1940)

On the non-existence of tests of 'students' hypothesis' having power function independent of o. *Annals of Mathematical Statistics,* 11, 186-192.

Davis, R. A. (1989)

Personality: Its use in selecting candidates for U.S. Air Force undergraduate pilot training (Research Rep. No. AU-ARI-88-8). Maxwell Air Force Base, AL: Air University Press.

Dawson, S. (1913)

Binocular and uniocular discrimination of brightness. *British Journal of Psychology,* 6, 78-108.

Debreu, G. (1960)

Topological methods in cardinal utility theory. In Karlin, S., & Suppes, P., (Eds.), *Mathematical Methods in the Social Sciences.* Stanford: Stanford University Press.

Dennis, J. E., & Woods, D. J. (1987)

New computing environments: *Microcomputers in Large-Scale Computing,* edited by A Wouk, Siam, 116-122.

Derman, C. (1957)

Non-parametric up-and-down experimentation. *Annals of Mathematical Statistics,* 28, 795-797.

De-Wet, D. R. (1963)

The roundabout. A rotary pursuit-test, and its investigation on prospective air-pilots. *Psycholgia Africana,* 10, 48-62.

Digman, J. M. (1989)

Development, stability and utility. *Journal of Personality,* 57 195-214.

Digman, J. M. (1990)

Personality structure: Emergence of the five factor model. *Annual Review of Psychology,* 41, 417-440.

Dixon, W. J., & Mood, A. (1948)

A method for obtaining and analyzing sensitivity data. *Journal of the American Statistical Association,* 43, 109-126.

Doignon, J. P., & Falmagne, J. C. (1974)

Difference measurement and simple scalability with restricted solvability. *Journal of Mathematical Psychology,* 11 (4), 473-499.

Doll, R. E. (1962)

Officer peer ratings as a predictor of failure to complete flight training (Special Report No. 62-2). Pensacola, FL: U.S. Naval Aviation Medical Center.

Driver, J., & Baylis, G. C. (1989)

Movement and visual attention: The spotlight metaphor breaks down. *Journal of Experimental Pyschology Human Perception and Performance*. 15,(3), 448-456.

Durlach, N. I., & Braida, L. D. (1969)

Intensity perception: I. Preliminary theory of intensity resolution. *Journal of the Accoustical Society of America,* 46, 372-383.

Dupac, V. (1984)

Stochastic approximation. In Krishnaiah, P. R., & Sen, P. K., (Eds.), *Handbook of Statistics,* 515-529. Elsevier.

Dwyer, J. H. (1983)

Statistical models for the social and behavioral sciences. New York: Oxford University Press.

Egan, J. P. (1975)

Signal Detection Theory and ROC Analysis. New York: Academic Press.

Ekman, G. (1956)

Subjective power functions and the method of fractionation. *Report from the Psychological Laboratory,* 34. Stockholm: University of Stockholm.

Ekman, G. (1961)

Methodological note on scales of gustatory intensity. *Scandinavian Journal of Psychology,* 2, 185-190.

Ekman, G. (1965)

Scaling. *Annual Review of Psychology,* 16, 451-474.

Ellis, B. (1966)

Basic Concepts of Measurement. London: Cambridge University Press.

Elshaw, C. C., & Lidderdale, I. G. (1982)

Flying selection in the Royal Air Force. Revue de Psychologie Appliquie, 32(Suppl.) 3-13.

Emerson, P. L. (1986a)

Observations on maximum likelihood and bayesian methods of forced choice sequential threshold estimation. *Perception & Psychophysics,* 39, 151-153.

Emerson, P. L. (1986b)

A quadrature method for bayesian sequential threshold estimation. *Perception & Psychophysics,* 39, 381-383.

Engen, T. (1971)

Psychophysics: Discrimination and detection. In Kling, J. W., & Riggs, L. A., (Eds.), *Experimental Psychology.* New York: Holt Rinehart & Winston.

Eriksen, C. W., & Yeh, Y. Y. (1985)

A location of attention in the visual field. *Journal of Experimental Pyschology Human Perception and Performance.* 11,(5), 583-597.

Fagot, R. F. (1963)

On the psychophysical law and estimation procedures in psychophysical scaling. *Psychometrika,* 28, 145-160.

Falmagne, J. C. (1968)

Note on a simple property of binary mixtures. *British Journal of Mathematical & Statistical Psychology,* 21 (1), 131-132.

Falmagne, J. C. (1971)

The generalised Fechner problem and discrimination. *Journal of Mathematical Psychology*, 8, 22-43.

Falmagne, J. C. (1974)

Foundations of Fexhnerian psychophysics. In Krantz, D. H., Atkinson, R. C., Luce, R. D., & Suppes, P., (Eds.), *Contemporary Developments in Mathematical Psychology*, 2, *Measurement Psychophysics and Neural Information Processing.* San Francisco: Freeman, W. H.

Falmagne, J. C. (1976)

Random conjoint measurement and loudness summation. *Psychological Review*, 83, 65-79.

Falmagne, J. C. (1977)

Weber's inequality and Fechner's problem. *Journal of Mathematical Psychology*, 16, 267-271.

Falmagne, J. C. (1978)

A representation theorem for finite random scales systems. *Journal of Mathematical Psychology*, 18, 52-72.

Falmagne, J. C. (1979)

On a class of probabilistic conjoint measurement models: Some diagnostic properties. *Journal of Mathematical Psychology*, 19, 73-88.

Falmagne, J. C. (1980)

A probabilistic theory of extensive measurement. *Journal of Philosophy of Science*, 47 (2), 277-296.

Falmagne, J. C. (1982)

Psychometric functions theory. *Journal of Mathematical Psychology*, 25(1), 1-50.

Falmagne, J. C. (1986)

Psychophysical measurement and theory. In Boff, K. R., Kaufman, L., & Thomas, J. P., (Eds.), *Handbook of Perception and Human Performance*. New York: John Wiley & Sons.

Falmagne, J. C., & Iverson, G. J. (1979)

Conjoint Weber laws and additivity. *Journal of Mathematical Psychology*, 20, 164-183.

Falmagne, J. C., & Iverson, G. J., & Marcovici, S. (1979)

Binaural loudness summation: Probabilistic theory and data. *Psychological Review*, 86, 25-43.

Falmagne, J. C., & Narens, L. (1983)

Scales and meaningfulness of quantitative laws. *Synthese*, 55 (3), 287-326.

Fechner, G. T. (1860)

Elements of Psychophysics. D. H. Howes and E. C. Boring. (Eds.), Rinehart and Winston.

Fernandez, C., & Goldberg, J. M. (1976)

Physiology of peripheral neurones innervating otolith organs in the Squirrel Monkey. Parts 1, 2, & 3 *Journal of Neurophysiology*, 39, 970-1008.

Findlay, J. M. (1978)

Estimates on probability functions: A more virulent PEST. *Perception and Psychophysics*, 23 (2), 181-185.

Finney, D. J. (1971)

Probit analysis. Cambridge: Cambridge University Press.

Fisher, R. A., & Tippett, L. H. C. (1928)

Limiting forms of the frequency distributions of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24, 180-190.

Fiske, D. W. (1947)

Validation of naval aviation cadet selection tests against training criteria. *Journal of Applied Psychology*, 31, 601-614.

Flanagan, J. C. (1947)

The aviation psychology program in the Army Air Forces (Army Air Forces Aviation Psychology Program Research Report No. 1) Washington, DC: U.S. Government Printing Office.

Fleischman, H. L., Ambler, R. K., Peterson, F. E., & Lane, N. E. (1966)

The relationship of five personality scales to success in naval aviation training (NAMI Rep. No. 968). Pensacola, FL: U.S. Naval Aviation Medical Center.

Fleishman, E. A. (1954)

Evaluations of psychomotor tests for pilot selection: The direction control and compensatory balance tests (Tech. Rep. No. AFPTRC-TR-54-131). Lackland Air Force base, TX: Air Force Personnel and Training Research Center.

Fleishman, E. A. (1956)

Psychomotor selection tests: Research and application in the United States Air Force, *Personnel Psychology*, 9, 449-467.

Flyer, E. S., & Bigbee, L. R. (1954)

The light plane as a pre-primary selection and training device: III. Analysis of selection data (Tech. Rep. No. AFPTRC-TR-54-125). Lackland Air Force Base, TX: Air Force Personnel and Training Research Center.

Fowler, B. (1981)

The aircraft landing test: An information processing approach to pilot selection. *Human Factors,* 23, 129-137.

Gescheider, G. A. (1976)

Psychophysics: Method and theory. Hillsdale, N. J.: Erlbaum.

Gigerenzer, G., & Strube, G. (1983)

Are there limits to binaural additivity of loudness? *Journal of Experimental Psychology: Human Perception and Performance,* 9(1), 126-136.

Gleick, J. (1988)

Chaos, Viking Press, Penguin.

Gnedenko, B. V. (1943)

Sur la distribution limite tu terme maximum d'une serie aleatoire. *Annals of Mathematics,* 44, 423-453.

Goebel, R. A., Baum, D. R., & Hagin, W. V. (1971)

Using a ground trainer in a job sample approach to predicting pilot performance (Tech. Rep. No. AFHRL-TR-71-50). Williams Air Force Base, AZ: Air Force Human Resources Laboratory, Flying Training Division.

Goldberg, J. M., & Fernandez, C. (1971)

Physiology of peripheral neurones innervating semicircular canals of the Squirrel Monkey. Parts, 1, 2 & 3. *Journal of Neurophysiology,* 34, 635-684.

Golding, J. F. (1993)

Perceptual scaling of whole body low frequency linear oscillatory motion. *Aviat. Space and Env. Med.,* 64, 636-640.

Gopher, D. (1982)

A selective attention test as a predictor of success in flight training. *Human Factors,* 24, 173-183.

Gopher, D., & Kahneman, D. (1971)

Individual differences in attention and the prediction of flight criteria. *Perceptual and Motor Skills,* 33, 1335-1342.

Gordon, T. (1949)

The airline pilot's jobs. *Journal of Applied Psychology,* 33, 122-131.

Graham, R. B. (1990)

Phsyiological Psychology. (1990). Wadsworth Publishing. Belmont, CA.

Gravetter, F., & Lockhead, G. R. (1973)

Criterial range as a frame of reference for stimulus judgments. *Psychological Review,* 80, 203-216.

Gray, D. E. (Ed.). (1957)

American Institute of Physics handbook. New York: McGraw-Hill.

Graybiel, A., & West, H. (1945)

The relationship between physical fitness and success in training of U.S. Naval flight students. *Journal of Aviation Medicine,* 16, 242-249.

Green, D. M. (1978)

An Introduction to Hearing. Hillsdale, N. J. Erlbaum.

Green, D. M., & Luce, R. D. (1974)

Variability of magnitude estimates: A timing theory analysis. *Perception and Psychophysics,* 15, 291-300.

Green, D. M., & Luce, R. D. (1975)
Parallel psychometric functions from a set of independent detectors. *Psychological Bulletin,* 82, 483-486.

Green, D. M., Luce, R. D., & Duncan, J. E. (1977)
Variability and sequential effects in magnitude production and estimation of auditory intensity. *Perception and Psychophysics,* 22, 450-456.

Green, D. M., & Swets, J. A. (1966)
Signal Detection Theory and Psychophysics. Los Altos, Calf.: Peninsula.

Green, D. M., & Swets, J. A. (1974)
Signal Detection Theory and Psychophysics. New York: Krieger.

Green, D. M., & Swets, J. A. (1990)
Stimulus selection in adaptive psychophysical procedures. *Journal of the Acoustical Society of America,* 87, 2662-2674.

Greene, R. R. (1947)
Studies in pilot selecton: II. The ability to perceive and react differentially to configuration changes as related to the piloting of light aircraft. *Psychological Monographs,* 61, 18-28.

Griffin, G. R., & McBride, D. K. (1986)
Multitask performance: predicting success in naval aviation primary flight training (Tech. Rep. No. NAMRL-1316). Pensacola, FL: U.S. Naval Aerospace Medical Research Laboratory.

Griffin, G. R., & Mosko, J. D. (1982)
Preliminary evaluation of two dichotic listening tasks as predictors of performance in naval aviation undergraduate pilot training (Tech. Rep. No. NAMRL-1287). Pensacola, FL: U.S. Naval Aerospace Medical Research Laboratory.

Guedry, F. E. (1974)

Psychophysics of vestibular sensation. In Kornhuber, H. H. (Ed.), *Handbook of Sensory Physiology*, Berlin: Springer-Verlag, 6ii:3-154.

Guilford, J. P. (1954)

Psychometric methods. New York: McGraw-Hill.

Guilford, J. P., & Lacey, J. I (1947)

Printed classification tests (Army Air Forces Aviation Psychology Program Research Report No. 5). Washington, DC: U.S. Government Printing Office.

Guinn, N., Vitola, B. M., & Leisey, S. A. (1976)

Background and interest measures as predictors of success in undergraduate pilot training (Tech. Rep. No. AFHRL-TR-76-9). Lackland Air Force Base, TX: Air Force Human Resources Laboratory, Personnel Research Division.

Gumbel, E. G. (1958)

Statistics of extremes. New York: Columbia University Press.

Gundry, A. J. (1978)

Thresholds of perception for periodic linear motion. *Aviat. Space Env. Med.*., 49, 679-686.

Haber, R. N., & Standing, L. (1969)

Direct measures of short term visual storage. *Quarterly Journal of Experimental Psychology,* 21, 43-54.

Hall, J. L. (1968)

Maximum-likelihood sequential procedure for estimation of psychometric functions. *Journal of the Accoustical Society of America,* 44, 370.

Hall, J. L. (1981)

Hybrid adaptive procedure for estimation of psychometric functions. *Journal of the Acoustical Society of America,* 69, 1763-1769.

Harvey, L. O. (1986)

Efficient estimation of sensory thresholds. *Behavior Research Methods, Instruments & Computers,* 18, 623-632.

Heimer, W. I., & Tatz, S. J. (1966)

Practice Effects, Knowledge or Results and Transfer in Pitch Discrimination. Technical report. Navtradevcen IH-52, Port Washington 1966.

Helmholtz, H. V. (1930)

Zahlen und Messen erkenntnis-theoretisch betrachtet, Philiosophische Aufsutze Eduard Zeller gewidmet, Leipzig, 1887. (Reprinted in Gesammelte Abhandl., 1895, 3 356-391). *Counting and measuring.* Princeton, N.J.: Van Nostrand.

Hesse, A. (1986)

Comparison of several psychophysical procedures with respect to threshold estimates, reproducibility and efficiency. *Acustica,* 59, 263-273.

Hertli, P. (1982)

The prediction of success in Army aviator training: A study of the warrant officer candidate selection process. Unpublished manuscript, U.S. Army Research Institute Field Unit, Fort Rucker, AL.

Holland, M. K., & Lockhead, G. R. (1968)

Sequential effects in absolute judgments of loudness. *Perception and Psychophysics,* 3, 409-414.

Holway, A. H., & Pratt, C. C. (1936)

The Weber ratio for intensive discrimination. *Psychological Review,* 43, 322-340.

Hosman, R. J. W., & Vaart Van, J. C. (1981)
Effects of visual and vestibular motion perception on control task performance. *Proceedings of the first European Annual Conference on Human Decision Making and Manual Control.* Delft University.

Hosman, R. J. W., & Vaart Van, J. C. (1981)
Effects of vestibular and visual motion perception on task performance. *Acta Psychologica*, 48, 271-287.

Hosman, R. J. W., & Vaart Van, J. C. (1983)
Accuracy of Visually Perceived Roll Angle and Roll Rate Using an Artificial Horizon and Peripheral Displays. Delf University of Technology, Department of Aerospace.

Hosman, R. J. W., & Vaart Van, J. C. (1984)
Accuracy of system step response roll magnitude estimation from central and peripheral visual displays and simulator cockpit motion. *Proceedings of the Twentieth Annual Conference on Manual Control.* NASA Ames research.

Hosman, R. J. W., & Vaart, J. C. Van (1990).
Motion perception and vehicle control. *Perception and Control of Self Motion.* (Eds.), Rik Warren and Alex Wertham, 145-170.

Huang, J., & Young, L. R. (1981)
Rotation about the vertical axis with a Fixed Visual Field. *Exp. Brain Res.*, 41,172-183.

Hunter, D. R. (1982)
Air Force pilot selection research. Paper presented at the 90th meeting of the American Psychological Association, Washington, DC.

Hunter, D. R. (1989)

Aviator selection. In M. F. Wiskoff, & G. Rampton (Eds.), *Military personnel measurement* (pp. 129-167). New York: Praeger.

Hunter, D. R., & Burke, E. F. (1990)

An annotated bibliography of the aircrew selection literature (Research Rep. No. 1575). Alexandria, VA: U.S. Army Research Institute.

Hunter, D. R., & Burke, E. F. (1994)

Predicting aircraft pilot-training success: A meta-analysis of published research. *International Journal of Aviation Psychology*, (4), 4, 297-313.

Hunter, D. R., & Thompson, N. A. (1978)

Pilot selection system development (Tech. Rep. No. AFHRL-TR-78-33). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, Personnel Research Division.

Hunter, J. E., & Schmidt, F. L. (1990a)

Dichotomization of continuous variables: The implications for meta-analysis: *Journal of Applied Psychology*, 75, 334-349.

Hunter, J. E., & Schmidt, F. L. (1990b)

Meta-analysis: Cumulating research findings across studies. Beverley Hills, CA: Sage.

Iverson, G. J. (1979)

Note: Conditions under which Thurstone Case III representations for binary choice probabilities are also Fechnerian. *Journal of Mathematical Psychology*, 20(3), 263-271.

Iverson, G. J. (1983)

Weber's inequality and asymptotic representations of binary choice probabilities. Submitted to *Journal of Mathematical Psychology*.

Iverson, G. J., & Pavel, M (1980)
Invariant properties of masking phenomena in psychoacoustics and their theoretical consequences. *SIAM-AMS Proceedings*, 13, 17-24.

Iverson, G. J., & Pavel, M. (1981)
On the functional form of partial masking functions in psychoacoustics. *Journal of Mathematical Psychology*, 24, 1-20. (a).

Iverson, G. J., & Pavel, M. (1981)
Invariant characteristics of partial masking: Implications for mathematical models. *Journal of the Acoustical Society of America*, 69, 1126-1131. (b).

Joaquin, J. B. (1980)
The Personality Research Form (PRF) and its utility in predicting undergraduate pilot training performance in the Canadian Forces (Working Paper No. 80-12). Willowdale, Ontario: Canadian Forces Personnel Applied Research Unit.

Kaernbach, C. (1991)
Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics*, 49, 227-229.

Kaplin, H. L. (1965)
Prediction of success in Army aviation training (Tech. Rep. No. 1142) Washington, DC: U.S. Army Personnel Research Office.

Kaplin, H. L. (1975)
The five distractors experiment: Exploring the critical band with contaminated white noise. *Journal of the Accoustical Society of America*, 58, 504-511.

Kershaw, C. D. (1985)
Statistical properties of staircase estimates from two interval forced choice experiments. *British Journal of Mathematical and Statistical Psychology*, 38, 35-43.

Kesten, H. (1958)

Accelerated stochastic approximation. *Annals of Mathematical Statistics,* 28, 41-59.

King, J. E. (1945)

Relation of aptitude tests to success of Negro trainees in elementary pilot training (Research Bulletin No. 45-52). Tuskegee Army Air Field, AL: Office of the Surgeon, Headquarters Army Air Forces Training Command.

King-Smith, P. E., Grisby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1991)

Evaluation of four different variations of the QUEST procedure for measuring thresholds. *Investigative Ophthalmology and Visual Science (Suppl.),* 32, 1267.

King-Smith, P. E., Grisby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994)

Comparison of the QUEST and related methods for measuring thresholds: Efficiency, bias and practical considerations. *Vision Research,* 34, 885-912.

Knight, S. (1978)

Validation of RAF pilot selection measures (Note No. 7/78). London: Ministry of Defence, Science 3 (Royal Air Force).

Koonce, J. M. (1981)

Validation of a proposed pilot trainee selection system. In R. S. Jensen (Ed.), *Proceedings of the First Symposium on Aviation Psychology* (Tech. Rep. No. APL-1-81, pp. 255-260). Columbus: Ohio State University, Aviation Psychology Laboratory.

Laming, D., & Marsh, D. (1988)

Some performance tests of QUEST on measurements of vibrotactile thresholds. *Perception & Psychophysics,* 44, 99-107.

Laming, D. R. J. (1983)
Sensory analysis. (technical report), Dept. of experimental psychology, Cambridge University.

Lane, G. G. (1947)
Studies in pilot selection: I. The prediction of success in learning to fly light aircraft. *Psychological Monographs,* 61, 1-17.

Leek, M. R., Hanna, T. E., & Marshall, L. (1992)
Estimation of psychometric functions from adaptive tracking procedures. *Perception & Psychophysics,* 51, 247-256.

LeMaster, W. D., & Gray, T. H. (1974)
Ground training devices in job sample approach to UPT selection and screening (Tech. Rep. No. AFHRL-TR-74-86). Williams Air Force Base, AZ: Air Force Human Resources Laboratory, Flying Training Division.

Levitt, H. L. (1970)
Transformed up-down methods in psychoacoustics. *Journal of the Accoustical Society of America,* 33, 467-476.

Lidderdale, I. G. (1976)
The primary flying grading trial interim report No. 2. RAF Brampton, England: Ministry of Defence, Royal Air Force, Headquarters Command Research Branch.

Lieberman, H. R., & Pentland, A. P. (1982)
Microcomputer-based estimation of psychophysical thresholds: The best PEST. *Behavior Research Methods, Instruments & Computers,* 14, 21-25.

Lindeman, H. H. (1969)
Studies on the Morphology of the sensory regions of the vestibular apparatus. *Ergebnisse der Anatomie und Entwicklungsgeschichte,* 42, 1-113.

Lindsay, P. H., & Norman, D. A. (1972)
Human Information Processing. Academic Press International Edition.

Luce, R. D. (1959)
Individual Choice Behavior. New York: Wiley.

Luce, R. D. (1963)
Detection and recognition. In Luce, R. D., Bush, R. R., & Galanter, E., (Eds.), *Handbook of Mathematical Psychology,* pp. 103-189. New York: Wiley.

Luce, R. D., & Green, D. M. (1974)
Neural Coding and psychophysical discrimination data. *Journal of the Accoustical Society of America,* 56, 1554-1564.

Luce, R. D., & Green, D. M. (1974)
The response ratio hypothesis for magnitude estimation. *Journal of Mathematical Psychology,* 11(1), 1-14.

MacAdam, D. L. (1942)
Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America,* 32, 247-274.

McFadden, H. B. (1941)
The permanence of effects of visual training on visual acuity. *Three Studies in Psychological Optics*. Duncan, Oklahoma: Optometric Extension Program.

McFadden, D., & Richter, M. K. (1970)
Revealed Stochastic Preferences. Department of Economics. University of California, Berkely.

McFadden, D., & Richter, M. K. (1971)
On the Extension of a Set Function to a Probability on the Boolean Algebra Generated by a Family of Events with Applications. Working paper 14, Department of Economics, University of California.

McGrevy, D. F., & Valentine, L. D. (1974)
Validation of two aircrew psychomotor tests (Tech. Rep. No. AFHRL-TR-74-4). Lackland Air Force Base, TX: Air Force Human Resources Laboratory, Personnel Research Division.

McKee, S. P., Klein, S. A., & Teller, D. Y. (1985)
Statistical properties of forced-choice psychometric functions: Implications of probit analysis. *Perception & Psychophysics,* 37, 286-298.

McRuer, D. T., Graham, D., Krendel, E. S., & Reisener, W. (1965)
Human Pilot Dynamics in Compensatory Systems. Theory, Models and Experiments with Controlled Element and Forcing Function Variations. (AFFDL-TR-65-15) Wright Patterson Air Force Base.

McRuer, D. T., & Krendel, E. S. (1974)
Mathematical Models of Human Pilot Behaviour. (AGARDograph No 188) NATO Advisory group for aerospace research and development.

Madigan, R., & Williams, D. (1987)
Maximum-likelihood psychometric procedures in two-alternative forced-choice: Evaluation and recommendations. *Perception & Psychophysics,* 42, 240-249.

Maloney, L. T. (1990)
Confidence intervals for the parameters of psychometric functions. *Perception & Psychophysics,* 47, 127-134.

Manski, C.F. (1977)

The structure of random utility models. In Eberlain, Kroeber-Reil Leinfellner and Schick, (Eds.), *Theory of decision.* Dordrecht, Holland.

Marschak, J. (1960)

Binary choice constraints on random utility indicators. In Arrow, Karlin, Suppes (Eds.), *Standard Symposium on Mathematical Methods in the Social Sciences.* Stanford University Press.

Martinussen, M., & Torjussen, T. (1993)

Does DMT (Defense Mechanism Test) only predict pilot performance in Scandinavia? In R. S. Jensen & D. Neumeister (Eds.), *Proceedings of the Seventh International Symposium on Aviation Psychology* (pp. 398-403). Columbus: Ohio State University.

Martz, H. F., & Waller, R. A. (1982)

Bayesian Reliability Analysis. New York: John Wiley & Sons.

Mead, A. D., & Drasgow, F. (1993)

Equivalence of computerized and paper and pencil congnitive ability tests: A meta-analysis. *Psychological Bulletin,* 114, 449-458.

Meiry, J. L. (1965)

The Vestibular System and Human Dynamic Space Orientation. MIT Man-Vehicle Control Laboratory Report. T-65-1. Cambridge Massachusetts.

Moray, N. (1959)

Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology,* 11, 56-60.

Nachmias, J. (1981)

On the psychometric function for contrast detection. *Visition Research,* 21, 215-223.

Neter, J., Wassermann, W., & Kutner, M. H. (1990)
Applied Linear Statistical Models. Boston: Irwin.

Norman, D. A. (1976)
Memory and Attention. An Introduction to Human Information Processing. Wiley and Sons.

Oppenheim, A. V., & Schafer, R. W. (1975).
Digital Signal Processing, Prentice Hall.

O'Regan, J. K., & Humbert, R. (1989)
Estimating psychometric functions in forced-choice situations: Significant biases found in threshold and slope estimations when small samples are used. *Perception & Psychophysics,* 45, 434-442.

Park, S. K., & Miller, K. W. (1988)
Random number generators, good ones are hard to find. *Communications of the ACM,* 31, 1192-1201.

Pelli, D. G. (1987a)
The ideal psychometric procedure. *Investigative Ophthalmology and Visual Science (Suppl.),* 28, 336.

Pelli, D. G. (1987b)
The ideal psychometric procedure. *Perception (Suppl.),* 16, 237.

Pentland, A. P. (1980)
Maximum Likelihood Estimation: The Best PEST. *Perception and Psychophysics.* 28 (4), 377-379.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992)
Numerical Recipes in C: the Art of Scientific Computing. Cambridge: Cambridge University Press.

Rabiner, L. R., & Gold, B. (1975)
Theory and Application of Digital Signal Processing, Prentice Hall,

Rawlings, E. I., Rawlings, I. L., Chen, C. S., & Yilk, M. D. (1972)
The facilitating effects of mental rehearsal in the acquisition of rotary pursuit tracking. *Psychonomic Science*, 26, 71-73.

Reich, J. G. (1992)
C Curve Fitting and Modelling for Scientists and Engineers. New York: McGraw-Hill.

Robbins, H., & Monro, S. (1951)
A stochastic approximation method. *Annals of Mathematical Statistics*, 22, 400-407.

Rose, R. M., Teller, D. Y., & Rendleman, P. (1970)
Statistical properties of staircase estimates. *Perception & Psychophysics*, 8, 199-204.

Sampson, A. R. (1988)
Stochastic approximation. In Kotz, S., & Johnson, N. L., (Eds.), *Encyclopedia of Statistical Sciences*. New York: Wiley.

Schiffman, H. R. (1976)
Sensation and Perception: An integrated approach. Toronto; Wiley.

Schmidt, F. L., Hunter, J. E., & Urry, V. (1976)
Statistical power in criterion-related validation studies. *Journal of Applied Psychology*, 61, 478-485.

Sen, P. K. (1985)
Theory and applications of sequential nonparametrics. Philadelphia: SIAM.

Shelton, D. R., & Scarrow, I. (1984)

Two-alternative versus three-alernative procedures for threshold estimation. *Perception & Pschophysics,* 35, 385-392.

Sheridan, Th. B., & Ferrell, W. R. (1974)

Man Machine Systems: Information Control and Decision. Models of Human Performance. Cambridge MA. MIT.

Shipley, E. (1961)

Dependence of successive judgments in detection tasks: Correctness of response. *Journal of the Acoustical Society of America,* 71, 1527-1533.

Siem, F. M., Carretta, T. R., & Mercatante, T. A. (1987)

Personality, attitudes, and pilot training performance: Preliminary analysis (Tech. Rep. No. AFHRL-TR-87-62). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Silberstein, L., & MacAdam, D. L. (1945)

The distribution of color matchings around a color center. *Journal of the Optical Society of America,* 35, 32-39.

Simpson, W. A. (1989)

The STEP method: A new adaptive psychophysical procedure. *Perception and Psychophysics,* 45, *572-576.*

Simpson, W. A. (1988)

The method of constant stimuli is efficient. *Perception and Psychophysics,* 44 (5), 433-436.

Smith, K. J. E. (1961)

Stimulus programming in psychophysics. *Psychometrika,* 26, 27-33.

Spackman, K. P. & Tan, M.S. (1993)
When the turning gets tough. *New Scientist*, 137, (1864), 26-31.

Spahr, J. (1975)
Optimization of the presentation pattern in automated static perimetry. *Vision Research,* 15, 1275-1281.

Sperling, G. (1959)
Information in a brief visual presentation. Doctoral Thesis, Harvard University.

Sperling, G. (1960)
The information available in brief visual presentations. *Psychological Monographs,* 74.

Spiegel, M. R. (1961)
Theory and Problems of Statistics, McGraw Hill International.

Stelmach, L. B., & Herdman, C. M. (1991)
Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human Perception and Performance*, 17, (2), 539-550.

Stillmann, J. A. (1989)
A comparison of three adaptive psychophysical procedures using inexperienced listeners. *Perception & Psychophysics,* 46, 345-350.

Stott, J. R. R., & Benson, A. J. (1990)
Vestibular tests on UK Astronaut Candidates for an Anglo/ Russian Space Flight (Juno Mission). *Proceedings of the Fourth European Symposium on Life Science Research in Space. (ESA SP-307, 1990).*

Suppes, P. (1963)
Basic Measurement theory. In *Handbook of Mathematical Psychology.* Zinnes, J.L. Vol 1, New York, Wiley.

Swanson, W. H., & Birch, E. E. (1992)
Extracting thresholds from noisy psychophysical data. *Perception and Psychophysics,* 51, 409-422.

Taylor, M. M. (1971)
On the efficiency of psychophysical measurement. *Journal of the Accoustical Society of America,* 49, 505-508.

Taylor, M. M., & Creelman, C. D. (1967)
PEST: Efficient Estimates on Probability Functions. *Journal of the Accoustical Society of America,* 41 (4), 782-787.

Taylor, M. M., Forbes, S. M., & Creelman, C. D. (1983)
PEST reduces bias in forced choice psychophysics. *Journal of the Accoustical Society of America, 74,* 1367-1374.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991)
Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology,* 44, 703-742.

Thorndike, R. L. (1949)
Personnel Selection: Test and Measurement Techniques. New York. Wiley.

Thurstone, L. L. (1927a)
A law of comparative judegment. *Psychophysical review,* 34, 273- 286.

Thurstone, L L. (1927b)
Psychophysical analysis. *American Journal of Psychology,* 38, 368 - 389.

Treisman, A. M. (1964)
Monitoring and storage of irrelevent messages in selective attention. *Journal of Verbal Learning and Verbal Behaviour,* 3, 449-459.

Treisman, A. M., & Rostron, A. B. (1972)
Brief Auditory Storage: A modification of Sperling's paradigm applied to audition. *Acta Psychologica*, 36, 161-170.

Treutwein, B. (1989)
Adaptive psychophysical procedures. *Perception (Suppl.),* 18, 554.

Treutwein, B. (1991)
Adaptive psychophysical methods. In Bhatkar, V. P., & Rege, K. M., (Eds.), *Frontiers in knowledge-based computing.* New Delhi: Narosa.

Treutwein, B., & Rentschler, I. (1992)
Double pulse resolution in the visual field: The influence of temporal stimulus characteristics. *Clinical Vision Sciences,* 7, 421-434.

Treutwein, B., Rentschler, I., & Caelli, T. M. (1989)
Perceptual spatial frequency-orientation surface: Psychophysics and line element theory. *Biological Cybernetics,* 60, 285-295.

Turnbull, G. J. (1992)
A review of military pilot selection. *Aviat. Space Environ. Med.*, 63, 825-830.

Tyrell, R. A., & Owens, D. A. (1988)
A rapid technique to assess the resting states of eyes and other threshold phenomena: The modified binary search (MOBS). *Behavior Research Methods, Instruments, & Computers,* 20, 137-141.

Volkman, A. W. (1858)
Ueber den Einfluss der Uebugn. Liebzig Berichte, *Math Phys., Classe*, 10, 38-69.

Wald, A. (1947)
Sequential Analysis. John Wiley and Sons, Inc, New York.

Wald, A. (1950)

Statistical Decision Functions. New York: Wiley.

Walsh, E. G. (1961)

Role of the vestibular apparatus in the perception of motion on a parallel swing. *J.Physiol.*, 155, 506-513.

Walsh, E. G. (1962)

The perception of rhythmically repeated linear motion in the horizontal plane. *Br J. Psychol.*, 53, 439-445.

Watson, A. B. (1979)

Probability summation over time. *Vision Research,* 19, 515-522.

Watson, A. B., & Fitzhugh, A. (1990)

The method of constant stimuli is inefficient. *Perception and Psychophysics,* 47, (1), 87-91.

Watson, A. B., & Pelli, D. G. (1979)

The QUEST staircase procedure. *Applied Vision Association Newsletter,* 14, 6-7.

Watson, A. B., & Pelli, D. G. (1983)

QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics,* 33, (2), 113-120.

Watt, R. J., & Andrews, D. P. (1981)

APE: Adaptive probit estimation of psychometric functions. *Current Psychological Reviews,* 1, 205-214.

Weibull, W. A, (1951)

Statistical distribution function of wide applicability. *Journal of Applied Mechanics,* 18, 292-297.

Weijer, C., & Elliot, C. (1996)

Down with placebolatry, *New Scientist*, 2019, 35.

Welford, A. T., (1980)

Reaction Times. London Academic Press.

Whitener, E. M. (1990)

Confusion of confidence intervals and credibility limits in meta-analysis. *Journal of Applied Psychology*, 75, 315-321.

Wilks, S. (1962)

Mathematical Statistics. New York: Wiley.

Woods, R. L., & Thomson, W. D. (1993)

A comparison of psychometric methods for measuring the contrast sensitivity of experienced observers. *Clinical Vision Sciences*, 8, 401-415.

Zacharias, G. L., & Young, L. R. (1981).

Influence of Combined visual and vestibular rotation. *Experimental Brain Research*, 41, 159-171.

Addendum

Figure 91 p158

Figure 91 was produced by normalising the Steer Integrals for each driving style so that the average Steer Integral in the Base condition was scaled to an arbitrary value of 100. The Steer Integrals at other tyre pressures were then scaled by the same scaling factor. This allows us to compare the % increase in metrics across different driving styles.
A similar normalisation process was used for Figures 94 - 96 p 161

Jorgensen Algorithm pp 116 - 118

a. The underlying function is a modified logistic function (which is a standard statistical function) modified so that the range lies within [0.5 1.0] to suit the 2AFC experimental design. The general formula for the logistic is;

$$1/Y = ab^x + g$$

b. The principle is to find the maximum likelihood fit of the modified logistic to the data.

c. The main advantage of using the logistic function is that we can determine the likelihood spaces of our estimations of the parameters a and b which define the threshold and slopes respectively. For example, Figure 77 shows the probability that we have determined the parameters a and b correctly for different combinations of a and b.
These parameters can be transformed to their corresponding thresholds and gradients as shown in Figures 75 and 76. The contour curves of Figure 76 show that with our choice of stimulus presentations that we have accurately determined the threshold at the expense of the slope of the graph (std dev).

d. A Monte Carlo simulation was conducted which compared the QUEST estimates of thresholds with the Jorgensen estimates. This simulation showed that the thresholds were equivalent but that the error estimates in the thresholds and slopes were more reliable for the Jorgensen algorithm than for the QUEST algorithm.