A Portfolio of Study, Practice and Research

Submitted for the
Doctorate of Psychology (PsychD)
in Clinical Psychology
Conversion Programme

University of Surrey

“An Investigation of the Role of Executive Functions in Cognitive Estimation Test (CET) Performance: implications of results from a task of inter-city distance estimation”

Drew H. Alcott

1999
## Contents of the Portfolio of Study, Practice and Research

**Drew H. Alcott**  
1999

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal Study Proposal</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Professional Dossier</strong></td>
<td></td>
</tr>
<tr>
<td>Curriculum Vitae: Post Qualification Experience</td>
<td>4</td>
</tr>
<tr>
<td>and Continuing Professional Development.</td>
<td></td>
</tr>
<tr>
<td>Professional Qualification: Clinical Psychology</td>
<td>11</td>
</tr>
<tr>
<td>MSc Degree</td>
<td></td>
</tr>
<tr>
<td>First Qualification: B.A. degree in Psychology</td>
<td>12</td>
</tr>
<tr>
<td>“Development of a Cognitive Rehabilitation</td>
<td>13</td>
</tr>
<tr>
<td>Programme for Adults with Acquired Brain Injury</td>
<td></td>
</tr>
<tr>
<td>at Unsted Park Rehabilitation Hospital”</td>
<td></td>
</tr>
<tr>
<td><strong>Academic Dossier</strong></td>
<td></td>
</tr>
<tr>
<td>“Confabulation: A Critical Review of Theories”</td>
<td>57</td>
</tr>
<tr>
<td>“A Review of the Clinical, Conceptual and</td>
<td>79</td>
</tr>
<tr>
<td>Empirical Adequacy of Measures of Awareness</td>
<td></td>
</tr>
<tr>
<td>Deficit Associated with Frontal Lobe Dysfunction”</td>
<td></td>
</tr>
<tr>
<td><strong>Research Dissertation</strong></td>
<td></td>
</tr>
<tr>
<td>“An Investigation of the Role of Executive</td>
<td>102</td>
</tr>
<tr>
<td>Functions in Cognitive Estimation Test (CET)</td>
<td></td>
</tr>
<tr>
<td>Performance : implications of results from a task</td>
<td></td>
</tr>
<tr>
<td>of inter-city distance estimation”</td>
<td></td>
</tr>
<tr>
<td><strong>Research Dissertation submitted for the MSc</strong></td>
<td></td>
</tr>
<tr>
<td>(Clinical Psychology)</td>
<td></td>
</tr>
<tr>
<td>“The Contribution of Image Change and Role</td>
<td>175</td>
</tr>
<tr>
<td>Fulfilment in the Present and Past to the</td>
<td></td>
</tr>
<tr>
<td>Psychological Well-being of Spouses Caring for</td>
<td></td>
</tr>
<tr>
<td>their Dementing Partner”</td>
<td></td>
</tr>
</tbody>
</table>
PERSONAL STUDY PROPOSAL
Psy.D. Clinical Psychology Conversion Programme

Name: Drew Alcott
Date: 9/10/96 (April 23, 1996)
Registration Date: April 1995
Registration Number: 3417638

1. CLINICAL DOSSIER

1.1 Clinical Training Since Qualification
Details of professional career work and progress, study, teaching and publication drawn from curriculum vitae.

1.2 Planned Work For Portfolio
Title: “Development of a Cognitive Rehabilitation Programme for Adults with Acquired Brain Injury at Unsted Park Hospital”

Summary of Issues
The background and context in which the programme was developed will be described. The programme that was developed will be described with reference to the Rusk Institute programme of Ben-Yishay. The differences, and reasons for them, between this and the programme at Unsted will be discussed. Finally, information that has been collated will be used to evaluate the programme to date.
2. ACADEMIC DOSSIER

2.1 CRITICAL REVIEW 1: Impaired Awareness after Brain Injury: Theory & Measurement of Unawareness

Summary of Issues
This essay will focus on the form of impaired awareness associated with frontal lobe damage. The main aim is to consider the adequacy of available measures for clinical and empirical use. Because measures should take into account the clinical phenomenon, its conceptualisation and theoretical accounts these will be summarised. The measures will be reviewed and their strengths and weakness will be discussed in this context.

2.2 CRITICAL REVIEW 2: Confabulation: A Critical Review of Theories

B.2. Summary of Issues
The theoretical accounts of confabulation will be reviewed and discussed with particular regard to the evidence on which they are based and the conclusions that have based on this evidence.

3. RESEARCH DOSSIER

3.1 Title: An Investigation of the Role of Executive Functions in Cognitive Estimation Test (CET) Performance: implications of results from a task of inter-city distance estimation

3.2 Research supervisor: Mr. Ian Davies (Professor in Psychology, UniS)
3.3 Background & relevance
The Cognitive Estimates Test (CET) has been used extensively in research and clinical practice since its publication in 1978 (Shallice and Evans). Although its theoretical meaning has been discussed it has not been explored empirically. Therefore the aim is to test the proposed explanation of Shallice and colleagues. The study will be of clinical as well as theoretical value as clinical interpretation.

3.4 Methodology
The estimation performance of people with and without brain injury will be compared. Several methods of measuring accuracy as well as measures of dysexecutive behaviour will be used.

Measures:
A new estimation test, the Milest Test, will be devised. This will allow the effect on accuracy of providing supplementary information to be compared between groups.

3.5 Data collection plan
Data will be collected from patients, primarily at Unsted Park, but also from colleagues were possible (e.g. at Ticehurst House Hospital, the Wolfson Centre).

Analysis:
Data will be analysed using SPSS. The main statistical method will be Analysis of Variance.

3.6 Help / resources needed
Computer statistical analysis, access to patients with a brain injury affecting frontal lobe functions and non-brain injured participants.

D. Alcott (1999) Personal Study Plan (revised); Submitted for the Doctorate in Clinical Psychology: University of Surrey, UK
# PROFESSIONAL DEVELOPMENT HISTORY

**Drew Alcott**

## EMPLOYMENT HISTORY

<table>
<thead>
<tr>
<th>Year</th>
<th>Position</th>
<th>Department</th>
</tr>
</thead>
<tbody>
<tr>
<td>1983</td>
<td>Completed Clinical Training</td>
<td>Exeter Heath Authority</td>
</tr>
<tr>
<td>1983-1986</td>
<td>Psychological Services for the Elderly (Basic Grade Clinical Psychologist)</td>
<td>Torbay Health Authority</td>
</tr>
</tbody>
</table>
| 1986-1988 | Psychological Services for the Elderly (7 sessions)  
Neuropsychological Services (3 sessions)  
(Principal Clinical Psychologist) | Coventry Health Authority          |
| 1988-1995 | Neuropsychological Services; community and general hospital (2 sessions)  
Neuropsychological Rehabilitation (8 sessions)  
(Principal Clinical Neuropsychologist progressing to Consultant Psychologist) | Weald of Kent Community Trust and  
Unsted Park Rehabilitation Hospital (concurrently) |
| 1995-1998 | Consultant Clinical Neuropsychologist; Head of Brain Injury Rehabilitation Clinical Programme (10 sessions) | Unsted Park Rehabilitation Hospital |
| 1998- | Consultant Clinical Neuropsychologist; Head of Brain Injury Rehabilitation Clinical Programme (8 sessions)  
Consultant Clinical Neuropsychologist (2 sessions) | Unsted Park Rehabilitation Hospital &  
Bournewood Community and Mental Health Trust, Chertsey |

D. Alcott (1999) Doctorate in Clinical Psychology; University of Surrey
### CONTINUING PROFESSIONAL DEVELOPMENT

**Conference & Training Seminar Attendance:**

**Pre-Doctorate Registration**

<table>
<thead>
<tr>
<th>Year</th>
<th>Event Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>“The Therapeutic Milieu Approach to Rehabilitation of the Brain Injured Person”, workshop by Dr. Ben-Yishay at Ticehurst House Hospital, Kent</td>
</tr>
<tr>
<td></td>
<td>“Purchasing Brain Injury Services in the NHS” conference at Hope Hospital, Salford</td>
</tr>
<tr>
<td></td>
<td>Visiting programme to services in the USA; Neuropsychology, Behavioural Health and Vocational Assessment and Rehabilitation (Eau Claire and Menomonie, Wisconsin and Harmarville, Pittsburgh)</td>
</tr>
<tr>
<td>1993</td>
<td>Study day on “Memory Rehabilitation; Theory and Practice” (Dr. N. Kapur and Dr. N. Moffat at Ticehurst House Hospital, Kent)</td>
</tr>
<tr>
<td></td>
<td>Two-day “Courtroom Skills Workshop” run by the BPS Special Group in Neuropsychology (London)</td>
</tr>
<tr>
<td>1992</td>
<td>Visiting programme to neuropsychologists and centres of Brain Injury Rehabilitation in the USA (New York, Denver, Minneapolis and the Mayo Clinic)</td>
</tr>
<tr>
<td></td>
<td>“Integrating Theory and Practice in Neuropsychology” workshop held by the BPS Neuropsychology SIG in Cambridge</td>
</tr>
<tr>
<td></td>
<td>Two-day conference on Neurorehabilitation at the National Hospital, Queens Square, London</td>
</tr>
<tr>
<td></td>
<td>“Perceptual and Cognitive Problems Following Brain Insult”, workshop by Prof. G. Humphries and Dr. J. Powell, Frenchay Hospital</td>
</tr>
<tr>
<td></td>
<td>Neuropsychology Special Interest Group conference on the Legal Aspects of Clinical Neuropsychology</td>
</tr>
<tr>
<td>1991</td>
<td>British Neuropsychological Society Conference at the National Hospital, Queens Square, London</td>
</tr>
<tr>
<td></td>
<td>Joint Experimental Psychology and Neuropsychology conference at Sussex University, Brighton</td>
</tr>
<tr>
<td>1990</td>
<td>Alzheimer Disease conference at the Institute of Psychiatry, London</td>
</tr>
<tr>
<td>1987</td>
<td>PSIGE Annual Conference in Edinburgh</td>
</tr>
<tr>
<td></td>
<td>Cognitive Therapy Annual Conference in Oxford</td>
</tr>
<tr>
<td>1986</td>
<td>Bereavement Training Workshop in Paignton</td>
</tr>
<tr>
<td></td>
<td>PSIGE Conference on Assessment of the Elderly in Bristol</td>
</tr>
<tr>
<td></td>
<td>BPS Annual Conference in London</td>
</tr>
<tr>
<td></td>
<td>Forensic Psychology Conference at Broadmoor Hospital</td>
</tr>
</tbody>
</table>

D. Alcott (1999) Doctorate in Clinical Psychology; University of Surrey
**PROFESSIONAL DEVELOPMENT (cont’d)**

<table>
<thead>
<tr>
<th>Pre-Doctorate Conference &amp; Training Seminar Attendance (cont’d)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1984-5 Sexual Therapy Course at Exeter District Psychology Department</td>
<td></td>
</tr>
<tr>
<td>1981-2 Institute of Family Therapy (London) Introductory Course</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conference &amp; Training Seminar Attendance: Post-Doctorate Registration</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1998 Workshop with Prof. Donald Stuss at Ticehurst House Hospital, Kent</td>
<td></td>
</tr>
<tr>
<td>“Executive Functions”. Conference of the British Neuropsychological Society (London)</td>
<td></td>
</tr>
<tr>
<td>Headway Annual Conference (two-day), Nottingham</td>
<td></td>
</tr>
<tr>
<td>Workshop with MacKay-Moore Sohlberg at Ticehurst House Hospital, Kent</td>
<td></td>
</tr>
<tr>
<td>“Malingering and Memory Assessment”; CCPCD at St. Thomas’s Hospital, London</td>
<td></td>
</tr>
<tr>
<td>1996 BPS Annual London Conference</td>
<td></td>
</tr>
<tr>
<td>“Driving After Brain Injury” conference at the annual Neuropsychology Special Group, Nottingham</td>
<td></td>
</tr>
<tr>
<td>1st World Congress in Neurological Rehabilitation, Newcastle upon Tyne</td>
<td></td>
</tr>
<tr>
<td>“Motivating the Brain Injured Patient” conference at the RNHD, Putney</td>
<td></td>
</tr>
<tr>
<td>1995 Experimental Psychology Society conference, Birmingham University Annual Conference of the BPS and the DCP, Warwick University</td>
<td></td>
</tr>
<tr>
<td>“Setting Standards of Practice in Head Injury Services” seminar in London</td>
<td></td>
</tr>
<tr>
<td>“Designing Head Injury Services to Meet the Needs of Patients”, seminar, Maidstone</td>
<td></td>
</tr>
<tr>
<td>“Psychotherapeutic Interventions with Brain Injured Patients”, workshop by Dr. G. Prigatano at Ticehurst House Hospital, Kent</td>
<td></td>
</tr>
<tr>
<td>Various papers presented during the year as part of the Neuroscience Forum at the University of Surrey</td>
<td></td>
</tr>
</tbody>
</table>

D. Alcott (1999) Doctorate in Clinical Psychology; University of Surrey
TEACHING and TRAINING

Throughout my career I have continued to provide training to a range of professionals including psychologists, doctors, nurses, social workers and other therapists (e.g. OT, PT, etc.). The post with Coventry Health Authority involved the most intensive training activities (see separate list). The core of these was a 5 day residential training course for qualified health care staff which was entitled “Working With Dementia”. The topics incorporated in this included;

- Introduction to the Dementias
- Managing Challenging, Problem Behaviours
- The Neuropsychology of Dementia
- Reality Orientation
- Normal Ageing
- Goal Planning
- Assessment of the Elderly
- Working with Carers
- Behavioural Analysis
- Community Services for the Elderly

These evolved into a variety of more in depth, higher level courses spanning 1-3 days in duration.

Teaching activities have included lecturing to postgraduate clinical psychologists in training on various courses at a variety of universities;
- The Southeast Thames Regional Training Scheme for Clinical Psychology
- University of Exeter
- University of Birmingham
- University of Warwick
- University of Surrey

The topics on which I have lectured include;

- Memory Disorders
- Frontal Lobe Disorders
- Anxiety Problems in the Elderly
- Neuropsychological Assessment
- Neuropsychology of Dementia
- Head Injury

I have also taught on the post-qualification training course organised by the BPS Special Group in Clinical Neuropsychology (London, October 23, 1998).

D. Alcott (1999) Doctorate in Clinical Psychology; University of Surrey
Conference Presentations:

Pre-Doctorate Registration:

None

Post-Doctorate Registration


RESEARCH and PUBLICATION

Pre-Doctorate Registration Publications


D. Alcott (1999) Doctorate in Clinical Psychology; University of Surrey
Pre-Doctorate Publications (cont’d)

Considering the FIM/FAM as a Measure for Evaluating Rehabilitation Progress; review of the literature and recommendations for its use at Unsted Park Rehabilitation Hospital. Alcott, D. (1994) (unpublished report; available from Unsted Park, Godalming, Surrey, GU7 1UW)

Outcome Measurement for Neurological Patients at Unsted Park Rehabilitation Hospital; developing disability rating scales and testing their validity. Alcott, D. and Conlon-Scott, M. (1994) (unpublished report; available from Unsted Park, Godalming, Surrey, GU7 1UW)

Post-Doctorate Registration Publications


D. Alcott (1999) Doctorate in Clinical Psychology; University of Surrey
Post-Doctorate Publications (cont’d)


AFFILIATIONS AND OTHER PROFESSIONAL ACTIVITIES

British Psychological Society (Associate Fellow)
Division of Clinical Psychology (Chartered Clinical Psychologist)
Special Group with an Interest in Neuropsychology (Full Member)
British Association of Behavioural and Cognitive Psychotherapy (Affiliate Member)
American Psychological Association (Full Member)

D. Alcott (1999) Doctorate in Clinical Psychology; University of Surrey
UNIVERSITY OF EXETER

We the undersigned hereby certify that

DREW HAROLD ALCOTT

having fulfilled the requirements of the
Ordinances and Regulations of the University
was by the authority of the Senate
admitted to the Degree of

MASTER OF SCIENCE

in Clinical Psychology

at a Congregation held in the University on

21 July 1986

[Signatures]

Chancellor

Vice-Chancellor

Academic Registrar and Secretary
UNIVERSITY OF READING

I hereby certify that

DREW HAROLD ALCOTT

having satisfied the Examiners in the Final Examination in

PSYCHOLOGY

and having in the Examination been awarded Honours in the First Division of the Second Class, has this day been duly admitted to the Degree of

BACHELOR OF ARTS

of this University

5 July 1980

Registrar
Development of a Cognitive Rehabilitation Programme
for Adults with Acquired Brain Injury
at Unsted Park Rehabilitation Hospital

Drew Alcott
(1999)

Submitted in part completion of the
Doctorate in Clinical Psychology at the University of Surrey.

Consultant Clinical Neuropsychologist
Head of Brain Injury Rehabilitation Clinical Programme
Unsted Park Rehabilitation Hospital, Godalming, Surrey
and
Bournewood Community and Mental Health NHS Trust
Abraham Cowley Unit, Chertsey, Surrey
PREFACE

This paper describes the programme which resulted from the involvement of myself and several other key members of staff at Unsted Park Hospital. Having worked at Unsted Park for 4 years by the time this project was begun I had developed various ideas which it had previously not been possible to pursue. Being involved in the project from the beginning provided the opportunity to bring many of these ideas to fruition. Although I have been able to influence the design of the programme I can not say that it has developed as a result solely of my activities or instructions. As in any team endeavour, all the staff are involved in the process of making decisions and taking action. I was fortunate that many of the ideas that I put forward have been translated into practice thanks to the assistance of the other team members. It is extremely difficult to isolate specific actions or occasions where the outcome might be attributed to what I had done. Therefore this type of project is quite different from the presentation of a single clinical case where the impact of the individual psychologist’s actions are usually clearer. The reader must appreciate that much of what is described in this paper represents the culmination of the author’s experiences, thoughts, plans and actions over several years. I do not claim that these are unique only to myself. There were several other key people involved. Of particular mention was Dr. Michael Oddy (Consultant Clinical Psychologist) whose direction and oversight as the Director of Head Injury Services, offered the benefit of years of experience with the patient group and experience of running a similar programme. Also crucial were the contributions made by the Team Leader, Carol Fowlie, Lisa Thomas (Speech & Language Therapist), Celia Leck (Nursing Sister), Nicki Antonevich and Fiona Addcock (Occupational Therapists). There have also been many other staff who have contributed since the programme was established.
Development of a Cognitive Rehabilitation Programme

for Adults with Acquired Brain Injury at Unsted Park Rehabilitation Hospital

Drew Alcott

CONTENTS

PREFACE 14
INTRODUCTION 16
BACKGROUND 16
The Beginnings 17
COGNITIVE REHABILITATION AT UNSTED PARK 18
Drawing Upon Current Models 21
Selecting Aspects of the NYU Approach Appropriate for Unsted Park 22
Description of the Rehabilitation Process 26
DATA COLLECTION AND ANALYSIS 30
Descriptive Information 30
Information Regarding Outcome 33
Discharge Placement and Support 39
Work status 39
FIM+FAM results 39
Neuropsychological Test Data 39
DISCUSSION 39
FUTURE POSSIBILITIES 46
REFERENCES 49
APPENDICES 52

Appendix 1 Core Principles
Appendix 2 Admission Criteria
Appendix 3 The Independence Continuum
Appendix 4 Referral Process
Appendix 5 The Group Programme
INTRODUCTION
This paper will describe the development and evaluation of a cognitive rehabilitation programme for adults with acquired brain injury. The development began in 1995 as an extension of the Brain Injury Rehabilitation Unit (BIRU) at Ticehurst House Hospital. My role as the full-time psychologist on site was to drive the process that was overseen by the Director of the Brain Injury Services, Dr. Michael Oddy. This initially involved establishing models and methods of practice and subsequently involved further developments to improve the programme.

In this paper I will describe the context in which the programme developed and discuss how this influenced the developments. The programme will be discussed in relation to the primary model which influenced its form. This will include consideration of the suitability of its various components to the current circumstances. I will also present some initial data regarding the participants and their progress.

BACKGROUND

Unsted Park Rehabilitation Hospital (UPRH) was a nursing home in the 1970's that gradually evolved into a rehabilitation hospital. This began by developing the physical therapies under medical supervision and latterly expanded to include a psychological service. As a result, a traditional model of rehabilitation
was adopted. Patients were admitted under the care of one of several medical consultants who met weekly with the therapists. Patients spent the day in their rooms from where they attended individual therapy sessions in the respective departmental bases.

Although efforts were made to develop an inter-disciplinary process this proved difficult for a variety of reasons. The variety of consultants meant several weekly “ward rounds” and differences in approach to rehabilitation. There was also considerable heterogeneity amongst the patients in terms of both diagnosis (e.g. stroke, spinal cord injury and Alzheimer’s Disease) and severity of disability. The therapists’ daily schedule was designed around providing each patient up to two sessions a day. Therefore it was very difficult arrange clinical meetings (e.g. goal planning) and to schedule group therapy sessions.

Although this type of programme is appropriate for many types of patients, it is neither appropriate or effective for those whose disabilities arise predominantly from cognitive impairments. For example, confused, mobile patients did not respond well when their programme contained extended periods without any structured activity. In cases where the physical rehabilitation was being hampered by behavioural complications (e.g. avoidance or non co-operation) it was not possible to achieve an appropriate, consistent approach under psychological direction.

The Beginnings

The development of cognitive rehabilitation at Unsted Park arose from an opportunity to expand existing services. The Brain Injury Rehabilitation Services were initially set up at Ticehurst House Hospital in 1989. There has since been an increasing awareness of the need for cognitive rehabilitation for some patients with acquired brain injured (e.g. see Tennant 1995, Cockburn & Gatherer 1988, Greenwood & McMillan 1993). This coupled with the success and growing reputation of the Ticehurst BIRU resulted in an increased demand.

Consequently I was invited to collaborate with the Director of the Ticehurst BIRU, Dr. Michael Oddy (Consultant Clinical Neuropsychologist) in the process of developing the service at Unsted Park.

After agreeing the core principles and admission criteria\(^1\) for the programme, staff were recruited internally in order to involve them from the earliest stage. The aim of this was to develop team cohesion as well as drawing upon the individuals' range of relevant experience. A series of meetings were then held with the core staff to develop the programme rationale and the therapeutic components of the programme (e.g. group therapies and goal setting meetings) and the team processes (e.g. weekly team clinical review meeting).

The time frame for completing the process and opening the service was relatively short. From the time when the general discussions were initiated in June 1995 the aim was to open the unit by October 1995. The actual time available to achieve this was also restricted due to the fact that the Director was not on site and I as well as the other staff continued to meet our clinical responsibilities in the existing neurological rehabilitation team. Despite this the programme did admit the first patients in the latter part of October 1995.

**COGNITIVE REHABILITATION AT UNSTED PARK**

There are several features that distinguish the programme from traditional neurological rehabilitation. Neurological rehabilitation focused upon physical problems (e.g. hemiparesis and walking) and is usually lead by a medical consultant whereas cognitive rehabilitation programme emphasises psychological problems (i.e. cognition, behaviour and emotion) and is lead by a neuropsychological consultant. This is appropriate given the nature of the problems being dealt with in this form of rehabilitation (see Oddy 1998).

\(^1\) See Appendices 1, “Core Principles” and 2, “Admission Criteria”.

The psychological focus is augmented by the use of Rehabilitation Assistants (RA’s). As psychology graduates, the RA’s are very capable and motivated to apply psychological principles. Their relevant knowledge is augmented by training from the clinical psychologist and other therapists. Because their practice is not influenced by previously established beliefs or practices, it is not necessary to modify inappropriate methods and attitudes.

In addition to the six Rehabilitation Assistants the staff includes an Occupational Therapist (OT) and Speech and Language Therapist (SLT). There is one full-time Neuropsychologist in addition to the Director, who is also a Consultant Clinical Neuropsychologist. There also is input from a Physiotherapist (PT) who sees the relatively few patients who require such treatment. The PT has also contributed to the Fitness Group. Medical cover was initially provided by a Neurologist but has been provided for most of the time by a Consultant in Rehabilitation Medicine.

The Team also includes nursing staff whose role is worth special mention. Because few of the patients require ‘nursing care’ (e.g. bathing, dressing) the role of rehabilitation nurse differs from the traditional nursing role. Because they have minimal physical care duties they are not required 24 hours a day. Therefore the rehabilitation nurse works a 12 hour day-shift (i.e. 08:00 - 20:00), with the late evening and night being covered by nurses from another neurological service in the hospital. This enables them to be present during the daytime when most patient’s goals are being addressed. In addition to some traditional nursing tasks (e.g. dispensing prescriptions), the rehabilitation nurse’s role includes generally over-seeing the patients’ needs during the day. They also over-see and supervise the rehabilitation assistants who are directly involved with the patients (e.g. prompting and supervising patients in their morning routine).

An analogy may help to describe the relative contributions of the different types of staff. Much of the content of a patient’s rehabilitation is provided by the qualified therapists. However, they have relatively limited contact with the patients (i.e. 1-2 hours a day per therapist). Therapists ‘create’ rehabilitation opportunities. They construct a situation that elicits the patient’s difficulties and then intervene. Rehabilitation Assistants have perhaps the greatest degree of contact with the patients and thus are in the best position to ‘capture’ rehabilitation opportunities. With sufficient training they become able to recognise when a patient’s difficulties are occurring during everyday tasks. With such training, direction and supervision from therapists the Rehabilitation Assistants become able to respond to this in a way that constitutes a rehabilitation intervention. Through their interventions the rehabilitation has greater intensity and continuity. Thus much of the substance of a patient’s rehabilitation is provided by the Rehabilitation Assistants. However, compared to the therapists, the Rehabilitation Assistants are relatively inexperienced. Therefore they need someone readily available to support and advise them. The rehabilitation nurse in the best position to do this. Therefore the bonding agent of a patient’s rehabilitation is provided by the rehabilitation nurse.

The programme is functionally focused aiming to increase independence in functional activities such as the routine of getting up in the morning, preparing meals and using time productively. Although the aims is to reduce handicap (WHO 1980, McGrath & Davis 1992) and improve quality of life, rehabilitation is not usually embarked upon unless incremental increase in independence is predicted. This can be conceptualised as a staircase of steps progressing from total dependency in hospital to full independence and productive use of time in the community2 (Alcott & Oddy 1998).

There is an underlying rationale that links the programme’s philosophy, principles and methods to the needs of the patients. Hence the design of group and individual therapies follows from the nature of the participants’ problems

---

2 See Appendix 3, “The Independence Continuum”.

and needs. For example, impaired social behaviour is common and an important determinant of outcome (see Ben-Yishay & Prigatano 1990). Therefore, the programme includes group therapies to address social behaviour and communication. The programmes structure and consistency are essential features in managing and reducing the effects of brain injury (e.g. poor memory, disorganisation, reduced initiation).

**Drawing Upon Current Models**

The design of the programme is also based on published knowledge and methods. Perhaps the most significant of these was the NYU rehabilitation programme directed by Dr. Yehuda Ben-Yishay. His approach has been adopted by numerous rehabilitation programmes in the UK as well as the USA.

During the early and mid-1970’s Ben-Yishay developed “the therapeutic milieu concept” (Ben-Yishay 1996). The original programme was developed for head injured military veterans of the 1973 Yom Kippur War in Israel. It was an intensive, lengthy, “no-expenses-spared” programme in which patients were involved for seven hours a day, six days a week, for a period of one year. The ratio of patients to staff was practically one to one. Ben-Yishay subsequently returned to the USA and developed a less costly, abbreviated form of this programme at the Rusk Institute, New York University (NYU) (see Ben-Yishay & Gold 1990). This retained some of the original’s key features. It is a 20 week out-patient programme run in a non-hospital setting for ten ‘trainees’ by six staff who are all clinical neuropsychologists (‘coaches’). It is a therapeutic milieu approach in which the closed group constitutes a ‘modified therapeutic community’. There is a consistent structure to each four-day week. During the five-hour long days there are various group therapy sessions in addition to the individual sessions with a specified member of staff who is the

---

3 For example, Ticehurst House Hospital, East Sussex and the Oliver Zangwill Centre, Cambridgeshire.

“trainee’s” coach. The programme combines cognitive remediation, psychotherapy and behavioural management.

Although Dr. Ben-Yishay’s programme has been particularly influential, it was neither possible or appropriate to simply replicate these methods. The local context (Unsted Park and Britain) differed in various ways from that for which Ben-Yishay developed his methods. Therefore key elements were selected and modified to suit the context.

Selecting Aspects of the NYU Approach Appropriate for Unsted Park

There are several core features of the NYU programme. It is an ‘holistic’ approach in which the ‘programmatics’ combine to form a ‘treatment continuum’. Holism refers to the way that the components of the therapy are interwoven (e.g. interdisciplinary rather than multidisciplinary). Programmatic concerns the manner in which the therapy is conducted, such as the structure of daily events and the way that staff interact with patients. The treatment continuum described by Ben-Yishay is a hierarchy of six overlapping phases. These are stages of psychological adjustment such as increased awareness of deficits, “malleability” (i.e. willingness to modify behaviour) and acceptance of the fact and consequences of a brain injury.

Some features of the NYU programme were not appropriate for the setting at Unsted. It was designed for a relatively small subset of brain injured patients (essentially those capable of returning to work). In my view, their programme is most suitable for patients at a later stage after they have had experience allowing them to test out their expectations (e.g. “I am back to normal now and able to resume work.”). Indeed Ben-Yishay has indicated that the programme is intended for patients who have previously completed a period of rehabilitation. In contrast, the programme at Unsted Park is primarily for patients earlier in the rehabilitation process when they are about to leave treatment in an acute hospital. Typically in this group the patient seeks

rehabilitation because they are aware that they are “not back to normal” but they expect to “make a full “recovery” including returning to previous activities such as work. At this stage they have had no experience that might challenge such expectations.

It would not be appropriate to run an out-patient programme at Unsted Park that was a ‘closed group’. It would be unsuitable for the more dependent patients such as those who are disoriented or not yet independent in self care. However, a relatively homogeneous group can be selected so that most share similar difficulties and due to the relatively long duration of rehabilitation the group membership is fairly stable.

A programme staffed by neuropsychologists, as at NYU, is appropriate when the main aim is psychological adjustment and adaptation. A multi-disciplinary approach is more appropriate for sub-acute rehabilitation where there are different aims. For example, occupational therapists have an important role in rehabilitation of self care and domestic skills. Nursing staff are necessary for patients requiring a degree of care or medical treatment (e.g. incontinence and anti-convulsant medication). Staff with specialist experience are essential to avoid misunderstanding patients’ behaviour and responding inappropriately.

Although the non-hospital setting advocated by Ben-Yishay has advantages this would not enable admission of some patients so early following initial hospitalisation. For example, some still require enteric feeding (e.g. PEG tube). Nevertheless a non-hospital atmosphere can be created (e.g. devan beds) and participants can be treated more like ‘trainees’ than ‘patients’.

Certain aspects of the NYU programmatics are unlikely to suit a British context. These include psycho-dramatic methods and the class-room atmosphere created by raising a hand before being allowed to speak during group sessions.

There were various core features of programmatics that were included. This includes the daily timetable. The timetable at UPRH was designed to be structured and consistent with the intention of helping to compensate for impairments of organisation and memory through its predictability. By avoiding extended periods of free time it also develops a routine to compensate for impaired drive and initiative. The extension of the timetable to include the entire waking day is a simple but important feature. Due to impaired awareness many patients need the implied message that they have difficulty with re-establishing a morning routine or using their leisure time productively. Without this they can feel that rehabilitation ‘finishes’ at five o’clock.

Programmatics also refers to the manner in which staff interact with the patients. The UPRH programme does not include the “dramatic” or “choreographic” elements of the NYU programme. However, the interactive manner has behavioural features. The approach of staff is guided by the principles of learning rather than being strictly behaviour modification. This includes staff responding to patients with an awareness of how they might positively reinforce appropriate behaviour. Likewise they are aware of how they might extinguish or avoid the development of inappropriate behaviour.

Principles of learning are also applied in the form of methods that facilitate learning. These include repetition (i.e. repeated practice) and the use of prompts and cues (i.e. discriminative stimuli). Maintaining the person’s motivation and participation is an important component of rehabilitation that also relates to these principles. By controlling the complexity of an activity (i.e. task difficulty) so that it is within the person’s current capacity it is more likely that they will succeed and thus reinforce the behaviours that are being re-learned. This is facilitated by adopting an “error-free” (Wilson 1996) approach to learning. Like the NYU programme, cognitive remediation is used to
increase awareness of deficits and then to encourage the development of compensatory strategies.

A therapeutic milieu that arises from an holistic approach is another feature of the NYU programme adopted at UPRH. More than one aspect of the programme is aimed at avoiding "artificially partitioning interventions". There is a degree of genericism amongst members of the multidisciplinary team due to some overlap in their roles and activities (e.g. all members function as keyworkers). ‘Partitioning’ can arise when patients’ goals are allocated according to profession (e.g. communication goals to the speech & language therapist, community goals to the occupational therapist etc.). This does not occur when responsibility for addressing goals is shared by more than one of the team members. In a less tangible way the philosophy and principles of the programme also contribute to the programme’s ‘holism’ as they link all aspects of the programme by affecting the design of group and individual interventions as well as all staff-patient interactions.

A “treatment continuum” is a part of the therapeutic milieu at UPRH although it is somewhat different from Ben-Yishay’s model. The continuum at NYU through which “the therapies are organised over several phases of the rehabilitation” relates to a model of psychological stages. This begins with ‘engagement’ (i.e. “optimising alertness, basic attention and concentration”) and strives towards “identity” (i.e. reintegration in to the community, at some level of productivity, with a “renewed sense of ego-identity”). As the rehabilitation at UPRH is primarily for patients in the early post-acute period the focus tends to be on the earlier stages of Ben-Yishay’s continuum (Ben-Yishay & Prigatano 1990). This is because their primary goals are ‘engagement’ and ‘awareness’. They are only just beginning to learn compensatory strategies involved in ‘mastery’ and ‘control’.

As mentioned above, a ‘continuum of independence’ has been useful in conceptualising and planning the rehabilitation at UPRH. Maximising independence in everyday activities is one of the primary aims of the UPRH programme (compared to the NYU aim of returning to work). The two aspects of this continuum relate to the complexity of the tasks in which the person is independent and the setting that this enables them to live within. Complexity ranges from the basic self-care tasks to high level productivity (e.g. employment). The setting ranges from hospital to independent accommodation in the community. At the lowest level the person is fully dependent and is in a hospital. They can progress through independence in self-care tasks, domestic activities and eventually paid employment. This would usually be related to progression through nursing home, residential home, supported residential accommodation and eventually independent accommodation. The value of this continuum is that it helps to identify rehabilitation objectives that are meaningful in terms of quality of life and costs to society.

Description of the Rehabilitation Process

Prior to admission all people referred are seen for a screening assessment. This is a fact-finding process that is discipline-free (Alcott & Oddy 1998). In addition to clarifying the nature of the problems and their cause (i.e. cognitive deficits due to acquired brain injury) the assessment aims to determine whether rehabilitation is likely to produce an incremental improvement in the situation.

Following the onset of participation there is a period of further assessment. Standardised measures are used to clarify the constellation of impaired and intact abilities. A measure of disability (i.e. the FIM+FAM; Hall 1992) is also completed. Together with clinical observations, this information is used to set the rehabilitation objectives and plan how they are to be addressed. The rehabilitation plan aims to be both realistic and functionally focused. In other words, the objectives are defined in terms of the various activities of living that
will significantly increase the person's degree of independence and quality of life. They are also set according to the predicted 'ceiling of recovery' and the amount of time likely to be available for the rehabilitation.

The rehabilitation programme combines group and individual therapies. The groups\(^4\) are designed to address the problems common to all the patients. It includes an orientation meeting in the morning during which the patients review the activities they will be doing that day and how these relate to their goals (and therefore implicitly their problems). The communication skills and cognitive skills groups provide opportunities to increase awareness and learn to compensate for impairments. The meal planning and meal preparation groups involve executive functions such as planning and organisation in the process of improving basic independence skills. It also involves behavioural skills such as co-operation and tolerance. There are educational discussion groups relating to general health (e.g. diet and the implications of a brain injury on driving and drinking alcohol) and understanding the effects of a brain injury. There are also group sessions to improve fitness and introduce various recreational activities as many people will have more 'leisure' time due to inability to resume work.

The individual therapy sessions serve several purposes. Each patient's particular constellation of difficulties can be addressed in an individualised way. They are able to begin learning methods to help compensate for impairments more easily in this setting. They also allow the clinicians to increase the person's awareness and acceptance of their difficulties more gradually and sensitively.

Goal planning meetings are held every three weeks where members of the team meet with the patient to discuss and agree the targets at which to aim during the subsequent weeks. Goals may be measurable targets or they may describe the means or methods of pursuing objectives. In addition to helping the patient

---

\(^4\) See Appendix 4, “The Group Programme”.

and team to remain focused upon the rehabilitation tasks the goals are also intended to facilitate collaboration (i.e. ‘engagement’; Ben-Yishay & Prigatano 1989). The nature of the goals depends upon the level of the person’s awareness deficit. Often the initial goals are fewer in number and more general in nature and as awareness increases they become more numerous and specific.

Various ways of ensuring good communication and team-work are needed as part of the process of monitoring progress and making required therapeutic adjustments. At a weekly “Clinical Review” meeting the team discusses each patient’s progress and any problems that have arisen. They also meet prior to each goal planning meeting to prepare what and how to introduce new, sometimes sensitive goals. A brief daily ‘handover’ meeting is important for communication and co-ordination. It enables rapid, immediate planning and decision making that can not await the weekly clinical review meeting.

For those aiming to return to work a phased approach is used. After engaging in work-type tasks in the hospital (e.g. catering, public speaking, paper-and-pencil problem solving tasks) the person will begin to return to their workplace while they continue to attend the programme. The Team advises and guides both the individual and the employer on various factors such as the length of the work day, the number of days per week and the type duties.

Preparation for discharge is an important part of the rehabilitation planning, hence it begins soon after admission. Although the degree eventual service input is not clear until later, the relevant agencies (e.g. social and/or employment services) are contacted early as it often takes considerable time to organise service provision. Because discharge is a major step for someone who has been in hospital for many months a gradual re-introduction to home and community is usually planned. This progresses from short visits home through weekends at home and then a phased reduction in the number of days attendance per week. We do not consider the individual’s rehabilitation to be
completed simply because they have been discharged. For people with moderate to severe brain injury rehabilitation should be seen as a lengthy period made up of various stages or phases. Increasing awareness and acceptance is a gradual process, often punctuated by "plateaux" during which they reject help. These too are phases in their rehabilitation. The process should not be considered complete until they reach the stage of 'Identity' (Ben-Yishay & Prigatano 1989) or reach the stage where further progress is unlikely. Therefore it is important to arrange follow-up monitoring and support. Prior to discharge further testing and assessment is done as part of outcome evaluation.

Previously outcome had been evaluated using a rating scale designed by UPRH. However, it lacked the sensitivity to detect change, particularly over the relatively short duration of rehabilitation (e.g. 3-4 weeks). In addition, it did not encompass the range of relevant disabilities. Consequently the clinicians had little faith in the measure and did not use it either for assessment and planning of the patients' rehabilitation or the evaluation of outcome. In the search for a new method the literature on evaluation measures was reviewed and another staff-designed measure was piloted (Alcott and Conlon-Scott 1994). Recommendations were subsequently accepted to adopt the use of the combined Functional Independence Measure and Functional Assessment Measure (FIM+FAM, Hall 1992; see Alcott 1994).

The FIM+FAM contains 30 items, each a seven point rating scale of independence/dependency. It is an amalgamation of the original 18 FIM items (Functional Independence Measure; Wright & Stone 1979, op.cit. Hall 1992) with the 12 additional FAM items (Functional Assessment Measure; Hall 1992). The items are divided into Motor and Cognitive scales, each of which contains several subscales. The subscales in the Motor scale are Self Care (Eating, Bathing, Grooming, Dressing, Toileting and Swallowing), Sphincter Control (Bladder and Bowel Management) and Mobility (Walking, Stairs, Community Access and Bed/Chair/Wheelchair, Car and Bath Transfers). In the
Cognition scale the subscales are *Communication* (Comprehension, Expression, Reading, Writing and Speech Intelligibility), *Psychosocial Adjustment* (Social Interaction, Emotional Status, Adjustment to Limitations and Employability) and *Cognitive Function* (Problem Solving, Memory, Orientation, Attention and Safety Judgement). Each item is a seven point rating scale reflecting the degree of dependency from totally dependent (1) to completely independent (7). The lower four points reflect different levels of assistance required (e.g. more/less than 50% assistance to perform the activity). The fifth point relates to supervision only while the upper two points refer to independent capability (complete or qualified, e.g. “concerns over safety”).

**DATA COLLECTION AND ANALYSIS**

The collection and analysis of data is considered to be an important activity that will help to understand the characteristics of the patients and how these may effect their response to rehabilitation and their longer term outcome. The data that has been collated concerns descriptive demographic information, duration of rehabilitation and measures of disability and impairment.

**Descriptive Information**

Descriptive information about the patients is shown in table 1. Typically there are 6-8 patients undergoing rehabilitation at any time. During the three years a total of 65 people have completed a period of rehabilitation, the majority of whom were NHS patients (73%). Most commonly referrals were made by a neuro-specialist (40%) or specialists in rehabilitation medicine (19%). Referrals were also made by other medical specialists (e.g. orthopaedics), general practitioners, psychologists, psychiatrists and solicitors. There were various causes of brain injury the most common being traumatic brain injury (TBI).
Table 1
Patient Descriptive Information

<table>
<thead>
<tr>
<th>Sex</th>
<th>Sex (%)</th>
<th>Age range (yr.)</th>
<th>Age mean (yr.)</th>
<th>Age (sd) (yr.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>82%</td>
<td>15-71 yr.</td>
<td>41 yr.</td>
<td>16 yr.</td>
</tr>
<tr>
<td>Women</td>
<td>18%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total=</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time since Onset (months)</th>
<th>Employment (OPCS code)</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1</td>
<td>1</td>
<td>12% TBI</td>
</tr>
<tr>
<td>1-2</td>
<td>2</td>
<td>11% SAH</td>
</tr>
<tr>
<td>2-3</td>
<td>3</td>
<td>15% Other CVA</td>
</tr>
<tr>
<td>3-6</td>
<td>4</td>
<td>4% Anoxia</td>
</tr>
<tr>
<td>6-9</td>
<td>5</td>
<td>4% Infection</td>
</tr>
<tr>
<td>9-12</td>
<td>6</td>
<td>5% Tumour</td>
</tr>
<tr>
<td>1-2 yr.</td>
<td>7</td>
<td>5%</td>
</tr>
<tr>
<td>&gt; 2 yr.</td>
<td>8</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>0-0</td>
<td>4%</td>
</tr>
</tbody>
</table>

Although the mean length of time since onset was five months there was considerable variation. About three quarters of the patients had suffered their brain injury within five months of commencement (see table 1).

Most of the patients were men. Considering that the incidence of head injury is greatest in young and older adults it is interesting that the age distribution was surprisingly even with two thirds falling in the range between 20-49 years. Although the majority were admitted from hospital (i.e. 62%) over a third were admitted from the community. The majority of patients were single at the time of admission (53%) with about one third (35%) married and the remainder either divorced or widowed (12%). Almost all had been working at the time of onset and they represented a wide range of occupations.

Typically patients had minimal physical impairments in combination with significant cognitive impairment. This is reflected in the results of the admission FIM+FAM ratings. Figure 1 shows the percentage of patients at three levels of FIM+FAM disability for each of the six subscales; Cognition (Cg),

5 Manager/Administrator (1), Professional (2), Associate Professional/Technical (3), Clerical/Secretarial (4), Craft & Related (5), Personal & Protective Service (6), Sales (7), Plant/Machine Operative (8), Other Manual (9), Student (0), Unemployed (0-0)

Psychosocial Adjustment (Ad), Communication (Cm), Mobility (Mo), Sphincter Control (Sp) and Self Care (SC). The three levels correspond to complete or near complete dependency (1-2), moderately dependent to supervision required (3-5) and independent (6-7). From figure 1 it is clear that the majority of the patients are independent on the Motor subscales (Mo, Sp and SC). On the Cognition sub scales (Cg, Ad and Cm) half or less were independent. A higher percentage were dependent and required assistance with these activities of daily life.

There was considerable variation in the duration of rehabilitation. Various factors affect the duration such as clinician’s initial prediction of the time that will be required and availability of funding. The mean duration has been 3 months (s.d.±1.93 months). The range of duration in illustrated in figure 2.
Information Regarding Outcome

Discharge Placement and Support

The type of accommodation into which the person is placed at discharge gives an indication of the level of independence and therefore it can be used to indicate outcome. Most of the patients (88%) were able to be discharged home. Equal proportions either went to residential accommodation or required further treatment (6% each). However, this is often only possible because they receive support. This was the case for about one in five (19%) who were supported either by a spouse or professional carer.

Slightly under a third (29%) of the people did not require any further assistance or support. A similar proportion (27%) received some level of support and about a quarter (23%) attended a specialist day centre (e.g. Headway). Quite a few people (20%) received support in the form of follow-up either from a case manager or health service clinician (e.g. neuropsychologist). A few people and/or their relatives declined the recommended services (5%). The details were unclear in about 10% of the cases.

Work status

Work status is one commonly used measure of outcome. Prior to entering the programme 10 of the patients were retired⁶ and therefore are not included in this analysis. At the time of discharge 73% were not working. Nine percent had returned to work part time, eleven percent were working full time and seven percent had returned to education. In other words, 27% had returned to work or education.

---

⁶ This includes six patients who had had their brain injury some time before and had since been retired on grounds of ill health.

FIM+FAM results

It is important to measure disability as an indicator of outcome considering that maximising independence is a central aim of rehabilitation. As mentioned, disability is assessed at admission and again at discharge using the FIM+FAM. Outcome can be assessed by comparing the results from these two occasions.

There are various methods used to complete the FIM+FAM (Adamovich 1992). These were considered prior to the implementation of FIM+FAM use at Unsted Park and a combined method was chosen (Alcott 1994). A “FIM+FAM co-ordinator” gathers and elicits relevant information from the rehabilitation team by asking pertinent questions (e.g. “Do they need any help to bathe?”, “Do they need help with more than 50% of the task?”) FIM+FAM derived from the decision making tree. The co-ordinator acts as arbiter to facilitate the process and avoid lengthy debate. This involves applying the basic rating principles laid down by the UDS (Uniform Data System) such as rating at the lowest observed level of performance. The co-ordinator who completed the rating form at admission is used to complete the form at discharge to minimise inter-rater variation.

Before reviewing the results of the FIM+FAM analysis it is worth considering the completeness of the data set. Full details (i.e. record forms completed at both admission and discharge) were available for 65 patients at admission and 58 patients at discharge. On some occasions one or more of the scales had not been completed. Due to its irrelevance in most cases, Wheelchair Mobility was the scale most commonly omitted. Excluding this scale only 3% of the scales were not rated at admission and only 1% were not rated at discharge. The Community Access/Mobility scale was the only other scale that was omitted other than on rare occasions. As it is difficult to quickly and easily assess this ability, it was not completed in 19% of the cases following admission. It was

---

7 UDS for Medical Rehabilitation, State University of New York at Buffalo, 232 Parker Hall, SUNY South Campus, 3435 Main Street, Buffalo, New York, 14214.
also incomplete in 5% of the cases at discharge. Therefore, the data set is sufficiently complete considering that only about 1% of the total potential data was unavailable.

There are however complications in using an ordinal scale, such as the FIM+FAM, to measure outcome (Alcott & Swann 1996, Alcott & Grafham 1996, 1997). Although the rating scale is numerical the resultant ratings should not be analysed as numbers. Consequently other methods of analysis must be found.

The method that has proven to be most useful is to determine the number of patients who meet or exceed a criterion of independence. The first criterion is derived from the FIM+FAM which attaches a rating of greater than four when the person can do the activity without assistance, although they may need prompting. Therefore the number of scales rated greater than four is first calculated. Determination of the second criteria is more arbitrary. At issue is the number of activities that the person can do independently such that they could be considered to be independent in a general sense. It would be unrealistic to expect them to be independent on all the scales (i.e. 100%). However, a reasonable but strict criteria would be that the person is independent on 95% of the scales. A less stringent criterion would be that they are independent on 75% of the scales.

![Figure 3](image)

**Figure 3**
Improvement in Physical and Cognitive Independence (with 95% criterion)

Figure 3 shows the difference between admission and discharge using the 95% criterion of independence. By the time of discharge there had been an increase in the number of patients meeting the 95% criterion on both Motor and Cognition scales. About 90% of the patients achieved this criterion on the Motor scale (compared to approximately 65% at admission). On the Cognition scale about 50% of the patients achieved the 95% criterion compared to less than 20% at admission.

Figure 4
Improvement in Physical and Cognitive Independence (with 75% criterion)

Figure 4 shows the difference between admission and discharge using the less stringent 75% criterion of independence. The same pattern can be seen of improvement on both Motor and Cognition scales with the latter improving to a greater degree.

As most of the patients have minimal physical disability at the time of admission it is not surprising that there is less improvement in this area and that about 90% are independent at the time of discharge. More importantly, there is a considerable increase in the level of independence in cognitive activities. About three quarters of the patients are 75% independent and about half are 95% independent in these activities of daily living.

Another indicator of change is the number of FIM+FAM scales on which the person has shown some improvement. As the largest and most valuable improvement relates to cognitive activities this is shown in figure 5. Seventy
percent of the patients show improvement on between one and five of the 12 cognitive items and 30% improve on more than 10 of these items.

![Figure 5](image)

Figure 5
Percentage of Patients With Improvement on Different Numbers of FIM+FAM Cognition Items

Improvement can also be illustrated in a more detailed analysis of the subscales. It can be seen in figure 6 that over 80% were independent on the subscales of Mobility, Sphincter Control and Self Care. More than 70% had achieved independence on the subscales of Communication and Cognition. However, fewer had achieved independence on the adjustment subscales compared to the communication and cognition subscales (i.e. approximately 50%). Nevertheless, the degree of change can be seen by comparing figure 6 with figure 1, where less than 50% were independent on Communication, less than 40% on Cognition and less than 25% on the Adjustment subscales.

![Figure 6](image)

Figure 6
Independence at Discharge:
Percentage of FIM+FAM subscale scores at 3 levels of independence (n=58).

Improvement was also analysed by comparing FIM+FAM scores at admission and discharge. The results of the Wilcoxon tests for matched samples is shown in table 2. It can be seen from this that there was a significant improvement on both scales and all subscales except Sphincter Control. This is unsurprising considering that nearly 90% were independent on this subscale at the time of admission.

Table 2

<table>
<thead>
<tr>
<th>Scales</th>
<th>(Max.)</th>
<th>Admission</th>
<th>Discharge</th>
<th>z-value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor</td>
<td>(105)</td>
<td>88.03</td>
<td>88.06</td>
<td>-3.46</td>
<td>.0005</td>
</tr>
<tr>
<td>Cognition</td>
<td>(98)</td>
<td>56.46</td>
<td>65.99</td>
<td>-4.31</td>
<td>&lt;.00009</td>
</tr>
<tr>
<td>Subscales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self Care</td>
<td>(49)</td>
<td>40.87</td>
<td>39.28</td>
<td>-1.85</td>
<td>.06</td>
</tr>
<tr>
<td>Sphincter</td>
<td>(14)</td>
<td>12.13</td>
<td>11.51</td>
<td>-0.04</td>
<td>.97</td>
</tr>
<tr>
<td>Mobility</td>
<td>(42)</td>
<td>35.03</td>
<td>37.27</td>
<td>-3.21</td>
<td>.001</td>
</tr>
<tr>
<td>Communication</td>
<td>(35)</td>
<td>24.25</td>
<td>25.42</td>
<td>-3.49</td>
<td>.0005</td>
</tr>
<tr>
<td>Adjustment</td>
<td>(28)</td>
<td>12.63</td>
<td>16.30</td>
<td>-4.19</td>
<td>&lt;.009</td>
</tr>
<tr>
<td>Cognitive</td>
<td>(35)</td>
<td>19.58</td>
<td>24.27</td>
<td>-4.02</td>
<td>.0001</td>
</tr>
</tbody>
</table>

Statistical significantly does not always correspond with clinical significance. The pre- versus post- rehabilitation difference on the Motor scale is statistically significant but there is no real change in the mean score. This finding is most likely due to the fact that one or two of the relatively few patients who had motor disabilities made some improvement over the course of rehabilitation. The most clinically significant improvement occurred on the Cognitive and the Psychosocial Adjustment subscales.

Neuropsychological Test Data

The results of neuropsychological tests can also be used as an indicator of outcome. However they are primarily an indicator of impairment and do not correlate strongly with disability or handicap. In addition, they tend to be relatively insensitive to change.

The tests routinely administered at present are shown in table 3. Unfortunately these could not be included as a substantial proportion of the results were unavailable at the time of writing this article. In many cases the difficulty was that standardised scores rather than raw scores had been stored hence it was not possible to complete an adequate analysis.

Table 3
Neuropsychological Assessments Routinely Administered

<table>
<thead>
<tr>
<th>TrailMaking Tests (TMT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revised Token Test (RTT)</td>
</tr>
<tr>
<td>Verbal Fluency (FAS)</td>
</tr>
<tr>
<td>National Adult Reading Test (NART)</td>
</tr>
<tr>
<td>Adult Memory &amp; Information Processing Battery (AMIPB)</td>
</tr>
<tr>
<td>Behavioural Assessment of Dysexecutive Syndrome (BADS)</td>
</tr>
<tr>
<td>Cognitive Estimates Test (CET)</td>
</tr>
<tr>
<td>Modified Wisconsin Card Sorting Test (M-WCST)</td>
</tr>
</tbody>
</table>

DISCUSSION

The process of setting up the programme was no doubt facilitated by the fact that it was a new development. It is likely that the process would have been more difficult had it been necessary to change an existing programme. The process was also facilitated by the good fortune of recruiting staff who shared a similar philosophy and set of ideas. Subsequent developments have been easier to achieve because of this positive foundation.

Ben-Yishay et al’s model from the NYU has been valuable, particularly in developing a psychological model of the aims and processes of rehabilitation. However, the psychological model has not yet been adequately studied. It is also difficult to empirically evaluate the programme and its components, particularly as its central tenets, the therapeutic milieu and holistic approach, are difficult to observe and therefore to measure. Nevertheless, there is evidence of positive outcomes for both the NYU programme and ours.
A temporary period during which various changes were made to the programme for organisational reasons created a 'natural experiment' to test our views on the importance of particular components of the programme. The programme was relocated, both geographically and organisationally. It was moved to another location in the building and 'merged' with another neurological rehabilitation programme. Although the changes that this brought about were not dramatic they were noticeable. The semi-closed nature of the programme was altered and there was a shift towards a more typical hospital environment. We noted that the participants spent more time in their bedrooms and less time in the social areas (e.g. the group room). It became more difficult to provide a milieu as the participants interacted with staff not trained or experienced in responding to such cognitive and behavioural changes. We were surprised at the extent of the changes and pleased to return to the original circumstances.

The emphasis on independence in everyday functioning, rather than performance on clinical tests, is particularly valuable. It makes rehabilitation more relevant to patients and therefore enhances collaboration. Its relevance is also more tangible to funders and is likely to have contributed to the degree to which they have supported the patients' continued rehabilitation. The use of the FIM+FAM contributes to this focus. However, it does not include many important ADL's (e.g. shopping and budgeting). The scope of ADL measurement as well as treatment needs to be extended. This is an issue that is being address through developing additional scales and expanding the facilities to encompass a wider range of ADL's in the rehabilitation process.
Another particular strength of the programme are the Rehabilitation Assistants. Their role is very important in ‘capturing rehabilitation opportunities’. Because they have a great deal of contact with the patients throughout the day in a variety situations they are able to prompt and reinforce desirable behaviour consistent with the goals. However, to be capable of this they need to be capable and psychologically oriented. This is achieved by selecting psychology graduates destined for clinical training. Regular staff turnover (e.g. to continue education or to gain further clinical experience) is a disadvantage. However, intelligent and motivated and learn their role and duties within a relatively short time.

All attempts to demonstrate that improvement is attributable to rehabilitation or specific features of the rehabilitation programme face an methodological challenge. Adequate comparisons are needed, such as a non-treatment control group or multiple baseline design. Although this paper has not set out to be a properly controlled outcome evaluation it is worth considering which features are worthy of study and how this might be done. Several features have been suggested as being particularly important to the efficacy and outcome of the rehabilitation. One was the structure and consistency of the weekly programme of therapy sessions (a.k.a. ‘programmatics’). Related to this is the provision of a daily timetable to facilitate orientation and insight. Another feature is the emphasis placed on independence in daily activities (i.e. reducing handicap through focus disability). The value of Rehabilitation Assistants with a psychological knowledge background was also mentioned. Specific components might also be studied. For example, the daily ‘Morning Orientation Meeting’ (MOM) is intended to improve awareness through the process of asking the patients to explicitly relate their sessions each day with particular rehabilitation goals.

The challenge involved in testing assertions about the value of such features is issue of how to control other factors to which improvement could also be
attributed. Spontaneous recovery is the most obvious factor. Possible methods to ‘control’ for this factor include starting rehabilitation once further spontaneous recovery is unlikely (e.g. after about 12 months) or using a non-treatment comparison group. There are ethical problems with delaying rehabilitation as there is evidence that early rehabilitation improves outcome (Cope 1995). Studying the importance of particular components of a rehabilitation programme is less ethically complicated as there is currently no clear evidence about these. Comparing existing programmes which differ (e.g. with versus without daily patient timetables) is one possible method. For example, within Unsted Park there are two neuro-rehabilitation programmes which could be compared. However, because there are a multitude of differences between the programmes a multi-factorial analysis would be needed. Other obvious factors to consider include severity of brain injury, time since brain injury and individual variables such as age and premorbid intellect. A more focused study would be another possibility. For example, using a measure of awareness it would possible to compare patients who had participated in MOM with those who had not (e.g. from the other rehabilitation programme).

The fact that the programme is provided by a private hospital raises interesting issues. Although it is an independent hospital in practical terms it is an NHS resource. The fact that the majority of patients are placed and funded by their NHS area attests to this. The relationship with the NHS has been excellent and mutually advantageous. From the clinician’s perspective their work is no different (i.e. their practice is driven by clinical, rather than business factors). In my opinion such a specialised service would be very difficult to provide by the NHS in its present structure. Because of the relatively low incidence of such patients it would not be economically viable for each area to operate a designated cognitive rehabilitation programme. In my experience it is not possible for a cognitive programme to coexist within another programme as it is practically impossible to create the necessary therapeutic milieu.
The evaluation of the programme has produced interesting findings but also revealed necessary improvements. This particularly concerns the collection of neuropsychological test data. Time will be needed to review many patients’ records to extract raw score results and collect missing data. The procedures have been improved by making data collection more comprehensive (e.g. including all or most tests) and by making collection an integral part of routine administrative procedure. Once sufficient neuropsychological test data is available it will possible to improve the description of patients at admission and discharge. It will also be possible to analyse relationships between neuropsychological variables and outcome (e.g. severity of dysexecutive syndrome and independence in ADL’s).

The FIM+FAM is not without some limitations. Although there is evidence supporting the reliability and validity of the FIM (e.g. Hamilton, Laughlin and Kayton 1991, Granger, Cotter, Hamilton, and Fielder 1994) there is minimal evidence published about the FIM+FAM. "Satisfactory" inter-rater reliability was reported by McPherson, Pentland, Cudmore and Prescott (1996) but this was qualified for several of the FAM items in the Cognition scale (i.e. Adjustment to Limitations, Speech Intelligibility, Orientation and Safety Judgement). They had previously reported that several of the FIM+FAM items were found by clinicians to be “difficult” to rate (Pentland & McPherson 1994). The suggestion that “FAM items involving abstract concepts such as ‘attention’ tend to be less reliable than directly observable behaviours” (Hall, Hamilton, Gordon & Zasler 1993) was supported by Alcott, Swann & Graffham (1997). They presented evidence that several of the cognitive items were indeed more abstract in nature. Limitations in the FIM+FAM’s reliability undermine its value in programme evaluation. It might lead to actual improvement being undetected or apparent improvement actually reflecting measurement error.
Although the FIM+FAM was intended to offer greater sensitivity to change than existing measures, such as the DRS (Hall 1992, Hall et al 1993) there is evidence that its sensitivity may not be sufficient. Hall, Mann, High, Wright, Kreutzer & Wood, D. (1996) reported significant ceiling effects in the FIM+FAM scales. By the time of discharge from rehabilitation more than 30% of the patients were scoring at or near the top of the scale (i.e. mean rating $\geq 6$) and at follow-up more than 70% were at ceiling. This has been noticeable in our experience, particularly with the Memory item. It only involves remembering or recognising people and commitments. After many weeks in the structured programme even people with severe memory impairment usually achieve this. Consequently by the time they are discharged they have reached the top of the scale. We have also found that the language items have an internal 'ceiling'. The increment between a rating of 5 (understands basic, non-abstract daily information) and 6 (understands abstract information) in practice is too great to detect improvement in many patients with non-dysphasia language impairments.

There is stronger evidence for the FIM+FAM’s validity. There is a strong correlation between the Motor and Cognition scales and another disability measure, the Disability Rating Scale (DRS) (Hall, Hamilton, Gordon & Zasler 1993). They also correlate significantly with measures of severity of injury (i.e. post traumatic amnesia) and measures of impairment such as the Ranchos Los Amigos Levels of Cognitive Functioning Scale (Hall et al 1993) and neuropsychological tests (McPherson, Berry & Pentland 1997). However, the study of McPherson et al (1997) did not find a clear relationship between the cognitive FIM+FAM items and their respective neuropsychological test (e.g. between the FIM+FAM ‘Attention’ item and tests of attention). Although we have found the FIM+FAM to be a valuable clinical tool, our experience has raised concerns about whether the cognitive items are valid measures of performance in cognitive activities of daily living (ADL’s) (see Oddy & Alcott 1996). For example, problem solving occurs in various different forms in

ADL's and we have felt that the existing Problem Solving item has limitations in reflecting this. The measure also lacks items for some ADL's that are crucial to independence and usually the focus of rehabilitation (e.g. shopping and cooking) consequently we have constructed scales for these in using the FIM+FAM format. The validity of the FIM+FAM is weakened for these reasons and therefore weakens its potential in programme evaluation.

The role of the psychologist in the programme is particularly interesting. In-patient programmes led by a clinical psychologist are still rare in Britain. Traditionally patients admitted to a centre registered as a hospital are admitted under a medic. Staff who are familiar with the traditional arrangement of medical leadership, particularly those whose own domain is more closely related to medicine (e.g. nursing, physiotherapy), may find this arrangement alien or uncomfortable. However, this has not been our experience. Staff are recruited who understand and accept the arrangement. In some cases the professional's role is different from that in physical rehabilitation. For example, the nursing role involves less physical 'care' and more psychological and behavioural intervention. The physiotherapy role involves a greater emphasis on fitness and relatively less traditional 'neuro-physio' treatment. To adjust to such differences the in hierarchy and role it is important that the staff feel their knowledge and skills are relevant, understood and valued. The leadership style that fosters this depends upon the clinician's personality rather whether they are a medic or a psychologist (Wood 1993).

Psychologists tend to work independently, as they have been trained to do, with perhaps some collaboration (e.g. colleagues in a family therapy team). Less common is the arrangement at UPRH and THH where senior psychologists work 'under' the directorship of another senior psychological colleague. This has the potential for disagreement over some decisions. For example, most senior psychologists would make their decisions about clinical matters (e.g. treatment approach) independently. They would also expect to make decisions
independently about psychological service development. Our experience has been quite the contrary. The arrangement has been a fruitful collaboration of mutual benefit. This too is likely to depend upon individual variables such as the degree of shared interests, practices and philosophy.

There are differences between my role as head of the clinical programme and that of the director of the service. Although both are consultants in terms of professional grade, the director's role is more that of an external consultant offering a broader experience and oversight. I act as the consultant to the team on psychological and neuropsychological matters on a day to day basis, whereas the director offers broader advice at weekly meetings. My role also includes provision of therapy which the director rarely does. The main exception to this that he runs a weekly 'Brain Injury' group that is an educational discussion forum for patients to talk about their personal experience of brain injury. I believe that direct, therapeutic contact with the patients prevents them from becoming too organisationally distanced and it contributes to their credibility in the team as well their understanding of the patients.

**FUTURE POSSIBILITIES**

The factors that influence future possibilities differ slightly between the public and private sector. 'Need' is a primary influence on development of public services, whereas 'demand' is relatively more important in the private sector. Although the need for brain injury rehabilitation services is likely to grow in the foreseeable future, the demand is likely to fluctuate. More people are surviving a brain injury due to improvements in acute medicine and changes in law the (e.g. helmets and seat belts), but the politics of health care will determine how this need is met. In the current political and financial situation there is a shift towards decreased use of non-public health services. This is likely to lead to a change in the type of patient as well as the number referred to private services.

It is important to evolve the service in line with the demand as well as producing evidence to substantiate assertions about the comparative benefits of different types of rehabilitation programme. If a cognitive rehabilitation programme, such as that at Unsted Park, has benefits not offered by other types of programme evidence is needed that demonstrates this and indicates which patients are most appropriate.

One possible direction for future programme development is towards the more ‘complex’, ‘challenging’ patients. Patients who are confused, agitated and mobile are relatively few in number but present a difficult challenge to most hospitals. They can ‘wander’ and so pose a safety risk to themselves which either requires a special, secure environment and/or intensive staff input. Brain injury can also caused marked irritability, sometimes with associated aggressiveness. This too is a challenging behaviour that most services are ill-equipped to manage and treat. Patients with multiple, severe impairments (e.g. spastic quadriplegia with orthopaedic injuries and impaired memory and awareness) require skilled and intensive nursing care and therapy. Their needs can exceed the resources of typical rehabilitation services. It is possible that there will be an increasing demand for specialised rehabilitation for these types of patients.

Another possible direction for programme development is to extend the current service. Supported residential accommodation is needed for a significant number of those who have suffered a moderate or severe brain injury. Most residential and nursing homes are designed for people who are older and often less capable of independence. As such homes are not appropriate for younger adults with a brain injury there is a need for accommodation that offers the right balance of support, assistance and independence. For some patients this type of setting can serve as ‘slow stream’ rehabilitation or as a ‘transitional living unit’ (TLU). A TLU component was recently set-up within the rehabilitation unit at Unsted Park with then intention of eventually expanding to

an off-site setting. Such accommodation could be linked with the rehabilitation unit in order to utilise the staffs’ skills and experience as a support resource.

Particular aspects of the rehabilitation programme could also be the focus for development. The transition process from inpatient rehabilitation to resettlement in the community could be developed. For example, a more systematic procedure for generalising compensatory strategies the patient has begun to use on the unit (e.g. mnemonic tools) might be developed. Improving methods of treating unawareness is another example of a potential clinical development.

Last, but not least, research is an important area for possible development. Research methodology and measurement tools are needed in order to study unawareness and develop related rehabilitation methods. A follow-up study is now in the initial stages of development. This will help to evaluate outcome, and to identify the need for further intervention or agency input.

REFERENCES


Alcott, D. (1994) Considering the FIM/FAM as a Measure for Evaluating Rehabilitation Progress; review of the literature and recommendations for its use at Unsted Park Rehabilitation Hospital. (unpublished report; available from Unsted Park, Godalming, Surrey, GU7 1UW)


Alcott, D. & Conlon-Scott, M. (1994) Outcome Measurement for Neurological Patients at Unsted Park Rehabilitation Hospital; developing disability rating scales and testing their validity. (unpublished report; available from Unsted Park, Godalming, Surrey, GU7 1UW)


Alcott, D. and Swann, R. (1996) Preliminary Results of the FIM/FAM in the Neurological Rehabilitation Services at Unsted Park; analysis of results, potential uses of the results and recommendations for action. (unpublished report; available from Unsted Park, Godalming, Surrey)


measured by the Functional Assessment Measure (FIM+FAM). 
*Neuropsychological Rehabilitation* 7, 241-257.


Appendix 1:

Core Principles

1. The patient must be assessed as being potentially able to benefit from rehabilitation (i.e. Be able to achieve/ work on Identifiable Goals)

2. Emphasis is placed on achieving improvement at a functional level (i.e. related to ‘real life’ activities).

3. Because all of the patients have suffered a permanent brain injury with associated impairments, on of the main purposed of the programme is to help them Form a New Identity -
   - With New Aspirations/Expectations
   - With New Sources of Satisfaction and Self Esteem
   - With New Ways to Use Their Time
   - With New Relationships
   - With New Abilities/Strengths

4. This corollary of this is the objective of helping patients learn to cope with their impairments through adjustment / compensating (i.e. to Use Strategies).

5. A structured environment is provided to help them cope and to facilitate learning.

6. A supportive environment is provided to maintain motivation and counteract depression and ‘Helplessness’ (i.e. grade the level of task difficulty to make success/achievement possible).

7. An environment which facilitates the learning process is also provided (e.g. Cueing, Errorless Learning, Precision Learning).

8. Patients are helped to establish ‘habits’/routine (i.e. Procedural Learning).

9. The rehabilitation process also includes work with family, relatives, employers, or future placements to extend progress achieved in the Unit into the home/community/work.

Appendix 2:

Admission Criteria

1) Have suffered a traumatic brain injury
(or other aetiology producing similar diffuse, non-progressive lesion).

2) Medically stable
(medical needs could be met at home if other rehabilitation needs were absent).

3) Age range 16-65 years.
(Some older patients may be considered suitable if the objectives for their rehabilitation are similar.)

4) Absence of congenital learning disabilities or major psychiatric history (prior to the brain injury),
drug of alcohol abuse.

5) The primary impairment will be cognitive, but there may be associated physical or behavioural
deficits resulting from the brain injury.

6) Potentially able to benefit from a period of rehabilitation.
Appendix 3:

THE INDEPENDENCE CONTINUUM

<table>
<thead>
<tr>
<th>Accommodation</th>
<th>Functional Activity Independence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent in own home in community</td>
<td>Independent at ‘work’ or supported ‘work’ Independent PUOT(^1) (leisure)</td>
</tr>
<tr>
<td>Own home with support input</td>
<td>Supervised/assisted PUOT (leisure) Semi-structured daily timetable and able to cope with longer ‘free’ periods</td>
</tr>
<tr>
<td>Staffed group residence</td>
<td>Structured timetable / routine Difficulty coping with extended periods of ‘free’ time Independent in domestic activities</td>
</tr>
<tr>
<td>Residential home</td>
<td>Not independent with domestic activities Independent in self-care</td>
</tr>
<tr>
<td>Nursing home</td>
<td>Assisted with self-care</td>
</tr>
<tr>
<td>Hospital</td>
<td>Dependent with self-care</td>
</tr>
</tbody>
</table>

\(^1\) PUOT - (Productive Use Of Time) The ability to engage in some form of activity that is enjoyable, fulfilling and purposeful.

Appendix:

Referral Process

Procedure for Admission to the Head Injury Programme

1) Initial enquiries may be made by telephone to the Head Injury Service. The Team Leader or Clinical Programme Director will discuss the potential referral with the enquirer.

2) In order to pursue the enquiry a letter of referral (usually from the Responsible Medical Officer) must be sent to the Director of Head Injury Services (Dr. M. Oddy).

3) The potential funnier will be contacted to inform them of the referral (if not already aware) and to have an initial discussion before further action is taken to follow-up the referral.

4) A senior member of the Team will see the patient to gather fuller information about their status and needs in order to determine whether the programme would be appropriate. This may be done through an outpatient appointment or a visit to see them in hospital. Some clinical details may be obtained by telephone.

5) The information which has been gathered will be presented at the weekly Team Meeting (Thursday) for discussion and decisions about:

   a) the suitability of the patient for the programme
   b) the possible of likely date for admission

6) This information is conveyed as appropriate to the referrer, funder, patient, and relatives.

7) Once funding is agreed the date for admission is fixed.

<table>
<thead>
<tr>
<th>Time</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:00-7:30</td>
<td>Morning Routine</td>
<td>Morning Routine</td>
<td>Morning Routine</td>
<td>Morning Routine</td>
<td>Morning Routine</td>
</tr>
<tr>
<td>7:30-8:00</td>
<td>Morning Meal</td>
<td>Morning Meal</td>
<td>Morning Meal</td>
<td>Morning Meal</td>
<td>Morning Meal</td>
</tr>
<tr>
<td>8:00-8:30</td>
<td>M.O.M.</td>
<td>M.O.M.</td>
<td>M.O.M.</td>
<td>M.O.M.</td>
<td>M.O.M.</td>
</tr>
<tr>
<td>8:30-9:00</td>
<td>Cognitive Group</td>
<td>Cognitive Group</td>
<td>Cognitive Group</td>
<td>Cognitive Group</td>
<td>Cognitive Group</td>
</tr>
<tr>
<td>9:00-9:30</td>
<td>M.O.M.</td>
<td>M.O.M.</td>
<td>M.O.M.</td>
<td>M.O.M.</td>
<td>M.O.M.</td>
</tr>
<tr>
<td>9:30-10:00</td>
<td>Mental Health Group</td>
<td>Commandeer Group</td>
<td>Commandeer Group</td>
<td>Commandeer Group</td>
<td>Commandeer Group</td>
</tr>
<tr>
<td>10:00-10:30</td>
<td>Meal Planning</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>10:30-11:00</td>
<td>Meal Preparation Group</td>
<td>Communication Group</td>
<td>Communication Group</td>
<td>Communication Group</td>
<td>Communication Group</td>
</tr>
<tr>
<td>11:00-11:30</td>
<td>Group Meeting</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>11:30-12:00</td>
<td>Fitness Group</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>12:00-12:30</td>
<td>R.A.P.</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>12:30-1:00</td>
<td>Private Time</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>1:00-1:30</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>1:30-2:00</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>2:00-2:30</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>2:30-3:00</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>3:00-3:30</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>3:30-4:00</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>4:00-4:30</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>4:30-5:00</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>5:00-5:30</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>5:30-6:00</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>6:00-6:30</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
<tr>
<td>6:30-7:00</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
<td>Mid-Day Meal</td>
</tr>
</tbody>
</table>

Notes & Plans

* M.O.M. = Morning Orientation Meeting
* R.A.P. = Review And Planning
* Cognitive Group = cognitive activities (e.g. memory, concentration, problem solving)
* SLT = Speech & Language Therapy
* OT = Occupational Therapy
* RA = Rehabilitation Assistant

Meal Serving Times:
- Breakfast (7:00-8:00)
- Mid-Day (12:00-13:30)
- Evening (5:45-6:30)
Confabulation: A Critical Review of Theories

Drew Alcott (1999)

INTRODUCTION

Confabulation is a clinical phenomena of memory impairment observed relatively often by clinical neuropsychologists. It is generally in the form of inaccurate recollections from autobiographical memory that go undetected by the patient, even when they are obvious to others (Moscovitch 1995; Burgess & Shallice 1996b). They may occur spontaneously or be elicited by questioning (i.e. 'momentary'; see Kopelman 1987) and range from credible to incredible (i.e. 'fantastic'; Berlyne 1972; see Kapur & Coughlan 1980). Although typically confabulations relate to episodic memory, they have been reported in semantic memories (Dalla Barba 1993b; Kopelman, Ng & Van Den Bourke 1997).

Research into confabulation has primarily involved studies of single cases (e.g. Kapur & Coughlan 1980, Dalla Barba 1993a,b) or groups of patients (e.g. Stuss, Alexander, Liberman & Levine 1978), although there have also been group comparisons (e.g. DeLuca & Cicerone 1991, Moscovitch & Melo 1997), and empirical studies. There are a variety of theoretical explanations including impairment of memory, monitoring of reality or memory source (Johnson 1991, 1993), strategic retrieval processes (Moscovitch 1995, Moscovitch & Melo 1997, Burgess & Shallice 1996a), output monitoring processes (Moscovitch & Melo 1997), frontal lobe dysfunction (Kapur & Coughlan 1980, Stuss, Alexander, Liberman & Levine 1978) and personality (Weinstien 1991).

Submitted in part completion of PsychD Clinical Psychology Conversion Course; University of Surrey, Guildford
The main aim of this paper is to form a foundation on which to develop ideas for further research. The literature will be reviewed to critically discuss relevant issues. The models and explanations that have been proposed will be reviewed to establish our current theoretical thinking and to consider what questions might be posed for further research. In concluding I will suggest research questions and recommend methods for exploring these.

**THEORETICAL ACCOUNTS OF CONFABULATION**

The model described by Burgess & Shallice (1996a) will be used to structure this review. They integrated several theories of memory with the aim of constructing a comprehensive model of confabulation. It is based on the distinction between associative retrieval, an automatic process driven by external cues, and strategic retrieval, a self-initiated process involving internal activation of cues to re-activate memories (Moscovitch 1989; op cit.). Strategic retrieval is seen as a three stage process involving retrieval specification (i.e. setting the criteria that describe the target memory), matching potential memory records with the target for evaluation (i.e. of the degree of fit between the target and located memory/memories) and modification (e.g. to account for new information that clarifies the target memory criteria) (Norman & Bobrow 1979; op cit.). Therefore, several aspects of retrieval could be involved in confabulation including the storage system, specification of the target memory, the retrieval process (i.e. "matching"), evaluation of potential memory output, monitoring of output and the effect of associative retrieval process output on strategic processes. The effect of individual differences (e.g. personality) will also be considered.
MEMORY FACTORS

Amnesia versus Impairment

Although confabulation has been attributed to amnesia (Berlyne 1972; cited in Fischer et al 1995) this can not be the cause as there are reports of individuals who confabulate but who are not amnesic (e.g. Kapur & Coughlan 1980; Papagno & Baddeley 1997). However, some impairment of memory is necessary. There are no reported cases of confabulation associated with normal scores on memory tests (Moscovitch & Melo 1997). Cunningham, Pliskin, Cassisi, Tsang & Rao (1997) found significantly greater memory impairment in the ‘high confabulation’ group than in the ‘low confabulation’ and ‘non-confabulation’ groups. The question is whether the impairment affects particular memory systems, memory processes or both.

Semantic versus Episodic

There are two questions regarding the issue of semantic memory confabulations; does confabulation affect semantic memory and if so what are the theoretical implications of this?

Although generally viewed as a disturbance of episodic memory, there is evidence of confabulation on semantic memory questions. Some of this evidence comes from single case studies. In a study comparing two patients, Dalla Barba (1993a,b) found that “confabulation was restricted to episodic memory tasks” in a patient with Binswanger’s Encephalopathy (BE) while it “affected the retrieval of semantic information” in a patient with a traumatic brain injury (TBI). Kopelman, Ng & Van Den Bourke (1997) also reported semantic memory confabulation dependent upon type of pathology. A patient with Alcoholic Korsakoff Syndrome confabulated across all seven question categories of the Confabulation Battery (Dalla Barba 1993a) including semantic memory items. A patient with frontal lobe damage also confabulated on some
of the semantic memory items while another patient with temporal lobe damage only confabulated on one of the 95 items.

Although clinical experience supports the empirical evidence of semantic memory confabulation these studies have limitations. The validity and reliability of Dalla Barba’s Confabulation Battery needs to be scrutinised. It is questionable whether some of the ‘Personal Semantic Memory’ items are truly semantic. As semantic memory refers to “knowledge about the world . . . and about words” (Tulving 1983), questions about changeable personal information are not truly semantic memory. This would include the items that ask about peoples’ ages (e.g. “How old are you?” and “How old are your children?”) as this is more likely to be ‘figured out’ rather than known (e.g. using the current year and birth date). Three of the items are arguably relate to episodic memory as they concern recent experiences (e.g. “When were you admitted to the hospital?” and “Have you seen me before?”).

There is evidence to support this concern about the Battery. There is some similarity between the performance of Kopelman et al’s control subjects and Dalla Barba’s TBI patient. Although Dalla Barba’s control subjects did not produce any confabulations on the 95 items, both of Kopelman et al’s confabulated on some of the items. In addition, the level of confabulation in these two was similar to the TBI patient. They confabulated on 33% of the General Semantic items compared to the patient’s 40%. They also confabulated on 20% of the Semantic and 20% of the Episodic ‘Don’t Know’ items compared to the patient’s 10% and 30% respectively. Consequently the total mean confabulation score for these categories of items was virtually the same for the controls as the patient (eight versus nine). Before further use as a measure of semantic confabulation, the Battery needs modification and validation.

---

2 Questions intended to elicit a 'don’t know' response, e.g. "What did you do on March 13, 1985?"
There is also evidence of semantic confabulation in a group comparison study by Moscovitch & Melo (1997). Having suggested that episodic confabulation was more prevalent in real-life conversations tasks such the Battery because they make greater demands on strategic retrieval, they used a cue-word task which they felt equalised the level of demand (e.g. "think of an experience that the word LETTER reminds you of"). In a group of patients with severe memory disorders they found proportionally more semantic than autobiographical confabulations in the confabulators group compared to the non-confabulator group.

This study is more soundly designed than most others. They assessed the veracity of answers to semantic items by referring to a reference source (e.g. encyclopaedia) and autobiographical answers by corroboration from a relative or testing consistency of response on more than one occasion. The criterion for categorising confabulators was also rational (i.e. >2 s.d. above the non-brain injured control group’s mean confabulation score). However, the cue-word task may not equalise retrieval demands on semantic and episodic memory. If it did we would not expect prompting to have a greater effect on retrieval from autobiographic memory as it did. If replicated it would interesting to compare the Confabulation Battery with the cue-word method. Based on Moscovitch's argument we would predict that the Battery would produce proportionally fewer semantic than episodic confabulations and that prompting would improve semantic memory retrieval more than on the cue-word task.

Evidence that confabulation can occur in either semantic or episodic form has led to different proposals regarding the theoretical implications. Dalla Barba (1993b) has argued that semantic and episodic confabulation reflect different forms of impaired monitoring processes based on two premises. Firstly, that episodic memory is a subsystem embedded within semantic memory (Tulving 1983). This leads to the proposal that semantic confabulation will occur when
the degree of memory impairment is more extensive. Secondly, that there are different "modes" of conscious awareness through which information is addressed, 'knowing' and 'remembering' (i.e. semantic and episodic memory). Hence because of the combination of 'degradation' of information in memory and impaired monitoring "the disruption of processes that normally monitor the retrieval of information is more widespread and also involves the 'knowing' mode of consciousness, resulting in confabulation in semantic memory". In Dalla Barba's view the existence of semantic confabulation is important because it indicates more extensive impairment affecting the semantic system and the 'knowing' mode of consciousness.

The evidence on which Dalla Barba bases his proposal comes from the study of two patients mentioned earlier. The confabulations of the patient with BE were said to be restricted to episodic memory while the TBI patient displayed semantic as well as episodic confabulations. The patients also responded differently in two experiments exploring the effect of cued recall under different encoding conditions. In cued recall the target word's category name was the cue (e.g. FRUIT- strawberry). During learning the words were presented alone the first experiment while in the second they are presented with the category name. The results are shown in figures 1 and 2

![Figure 1](image-url)  
Recall of Single Unrelated Words

3 Approximated from the actual figures in Dalla Barba (1993b).
Although both patients made dramatically more intrusion errors during cued recall compared to free recall, the patient with BE showed a less dramatic increase when encoding was enhanced in the second experiment. Dalla Barba interpreted the intrusions as 'confabulatory-like behaviour' and argued that the encoding condition in the first experiment “taps episodic memory, since it is crucial to remember the temporo-spatial context in which the items were acquired”. The encoding condition in the second experiment is one “in which also semantic memory is likely to be involved, in the sense that the retrieval is likely to be facilitated by the encoding of a semantic association”. Hence he infers that semantic memory monitoring processes were intact in the BE patient but impaired in the TBI patient. He also suggests that the virtual absence of intrusions during free recall is either because monitoring processes are functioning normally or are not involved.

The suggestions made by Dalla Barba to account for the absence of intrusions during free recall are not supported by any evidence and are unbelievable. Not only is it unlikely that monitoring processes would operate normally under one condition when they do not in the other, but in addition, the evidence of confabulation occurring without prompts suggests that monitoring is impaired (see Moscovitch & Melo 1997). It is also questionable whether the encoding

condition in the first experiment does not involve semantic memory as the presentation of words will automatically activate associated meaning. His conclusion would be more viable if the items to be recalled had no inherent meaning (e.g. non-words). In that case the two conditions would be distinguishable as the cue word in the second experiment would produce associated meaning that did not exist when the item appeared on its own.

An alternative interpretation of the differential effect of the encoding condition on cued recall is that it reflects differences in the patients' learning abilities. Because the percentage of intrusion errors and correct responses is less than 100% we can estimate the percentage of 'don't know' responses. This indicates that they are similar for each patient on the first and second experiment. Therefore, because the level of intrusions is the reciprocal of the number correct what has happened is that the BE patient has got more correct in the second than in the first experiment. This suggests that during acquisition he benefited from the cue words while the TBI patient did not. Consequently the patients differ between in terms of learning rather than retrieval and this does not indicate a different ability to monitor the output from semantic memory.

Kopelman et al. (1997) in the article mentioned earlier argue that their single case study does not support Dalla Barba's view. They concluded that although there was evidence of general semantic deficits and there was semantic confabulation, her confabulations were of "normal semantic structure" and were not 'fantastic'. This is contrary to Dalla Barba's argument that semantically anomalous or fantastic confabulations will occur as a result of semantic deficits. However, Kopelman et al.'s evidence of semantic deficit is weakened through being based on tests depending more on lexical than semantic processing (i.e. the NART & SCOLP). One of the SCOLP subtests does however, involve semantic processing and her low verbal IQ score on the

---

4 Number of 'don't know' responses = 100% - (% correct + % intrusions)
5 National Adult Reading Test and the Speed of Comprehension and Language Processing.
WAIS-R does support the suggested semantic deficit. Unfortunately they did not report details of WAIS-R subtests that would add support (e.g. Vocabulary, Similarities). There is also the issue, discussed earlier, of whether the Confabulation Battery reliably indicates semantic confabulations.

The theoretical significance of semantic confabulation is seen by Moscovitch & Melo (1997) as having implications for the integrity of retrieval processes. They argue that strategic retrieval is needed to recall memories with a “narrative structure”, such as episodic memories. As the structure of semantic memories is “based on other principles” they do not involve strategic retrieval and hence are less likely to evoke confabulation. They argue that semantic confabulation occurred in their cue-word study because the task involved narratives and therefore required strategic retrieval.

The relevant evidence is the similar frequency of confabulation for historical and personal (i.e. ‘semantic’ and ‘episodic’) cue words. There are also two assumptions that are relevant. One is that the cue-word method equalises the level of demands on semantic and episodic retrieval processes, and by implication that other’s methods did not (see also Moscovitch 1995). The second is that their recall task is ‘narrative’. In addition, they mention a significant correlation between the scores on the historical and personal conditions which they suggest indicates that “common factors” are involved in the retrieval mechanisms involved. However, their conclusions are weakened by assumptions that are not supported by evidence. They assume that the cue-word method equalises strategic retrieval demands on semantic as well as episodic memory while the Confabulation Battery does not. However, the greater impact of prompts on semantic items mentioned above does not support this. Secondly, they did not compare the cue-words with the Battery items to determine whether the level of semantic confabulation actually differs between these methods.

Retrieval Processes

The model of retrieval that can best account for much of the research evidence is Tulving's (1983) synergistic ecphory model that distinguishes two types of retrieval; strategic and associative (Baddeley 1997). It is strategic, rather than associative retrieval that is implicated in confabulation and is said to "cause the major positive and negative signs of confabulation" (Moscovitch & Melo 1997). The fact that confabulation affects premorbid memories, presumably stored when encoding and consolidation processes were intact implies that they are not involved (Moscovitch 1995). In addition, confabulating patients' recall improves when retrieval is facilitated, for example by providing cues (Dalla Barba 1993b, Moscovitch & Melo 1997) which indicates impaired ability to generate the cues needed to initiate memory search (Dalla Barba 1993b).

Retrieval involves both a mechanism to initiate the search and a mechanism that specifies the features of the target memory (Burgess & Shallice 1996a). The question is whether confabulation involves one or both of these. The cueing methods used by Dalla Barba (1993b) and Moscovitch & Melo (1997) do not clarified this issue as they potentially initiate a search as well as specifying the target memory characteristics. In Dalla Barba's experiments the cue words indicate the semantic features of the target (i.e. the word category). In Moscovitch & Melo's experiment the prompts specify various features of the target such as contextual (e.g. "Do you remember who was present?") and temporal (e.g. "What time of the day or in what year did it occur?"). In both studies providing a prompt potentially activates memory search.

Comparison of Dalla Barba's first and second experiments might be interpreted as demonstrating a dissociation between search initiation and target specification. Impaired search initiation in both patients is suggested by the finding in the first experiment that both produced significantly more responses (correct plus erroneous) when recall was cued. However, the TBI patient did
not increase the number of correct responses to the same extent as the BE patient in the second experiment when the cue-word was associated with the target word during acquisition. Because the cue-word in this experiment specifies the target as well as initiating search it indicates that the TBI patient, who benefited less, also has impaired target specification. This suggests impaired search initiation contributes to both forms of confabulation but spontaneous, rather than provoked confabulation is associated with impaired target specification.

Because Dalla Barba’s sample size is small, a group comparison study would be worthwhile. His second experiment could be modified to use the cue specificity effect (i.e. cues facilitate recall more if they have been associated with the target word during acquisition) (e.g. Tulving & Osler 1968 cited in Baddeley 1997). If half the cues during retrieval were new (e.g. colour-ORANGE during acquisition and fruit-ORANGE during recall) impaired initiation would produce depressed free recall but normal performance on both new and old cues. Impaired target specification would produce normal performance for old cues but depressed performance for new cues.

There have been no studies to determine which target specification features contribute to or ameliorate confabulation. Assuming that retrieval involves the reinstatement of the same pattern of cognitive activity that occurred at the time of encoding (Craik, Naveh-Benjamin & Anderson 1998) there is likely to be a wide variety of patterns across the spectrum of episodic memories. As Craik suggests that the control processes are different for encoding and retrieval, research might explore whether either or both of these improve as confabulation resolves. The divided attention tasks that Craik et al used to dissociate these control processes might prove useful in this.
Impaired search initiation and target specification alone can not fully account for confabulation. Impairment of strategic retrieval processes, such as target specification, does not always result in confabulation. Hanley, Davies, Downes & Mayes (1994) presented evidence of persisting strategic retrieval impairment causing severely impaired memory in a patient who had ceased to confabulate. The fact that prompts and cues increase confabulation and numbers of intrusions (e.g. Dalla Barba 1993b; Moscovitch & Melo 1997) indicates that facilitating target specification does not only increase recall of the correct memories. This could indicate impaired control of response output (i.e. failure to inhibit inappropriate responses) such as that found with frontal lobe dysfunction (Burgess & Shallice 1996b). Therefore other factors must contribute to confabulation.

NON-MEMORY FACTORS

Faulty comparator
A ‘mediator’ process that checks whether the potential memory fits the desired specification was included in Burgess & Shallice’s model (1996a). Impaired ‘reality monitoring’ was suggested by Johnson (1991) to account for memory errors such as confabulation. This is the process of discriminating the source of memories such as internal (e.g. imagined) from external (i.e. experiential perceptions). Johnson’s Multiple-Entry Modular system (MEMORY) contains a ‘reflective’ system, one level of which is ‘executive’ processes. This is involved in retrieving memory records similar to strategic retrieval (Johnson & Hirst 1993). This includes identifying the source of a record. Johnson places particular significance on being able to distinguish between external and internal sources of which there may be various (e.g. heard or read; thought, intended or dreamt). When reality monitoring fails and the source of a record is confused or misattributed confabulation can occur.

In support of the MEM model, Johnson, draws on a wide range of relevant literature as well as research by herself and colleagues (Johnson 1991, Johnson, Hashtroudi & Lindsay 1993, Johnson & Hirst 1993). There is evidence that people do distinguish between internally and externally generated information (Johnson & Raye 1981). Information that is self generated (i.e. internal) tends to show an advantage in recall and recognition tasks. Johnson has also demonstrated that internally generated information (e.g. imagined) can interfere with recall and recognition. This 'confusion effect' is more pronounced in 'good imaginers' (i.e. people who generate more vivid imaginations).

A patient said to be “no longer able to discriminate representations of wishes and imaginings from memories of actual experience” studied by Conway & Tacchi (1996) might also provided support for Johnson’s account. Following a head injury causing bilateral temporal and frontal lobe damage her confabulations appeared to be based on wishes that would offer a less painful account of traumatic experiences. Unfortunately, no evidence other than their interpretation of the confabulations were provided to support the argument that these represented wishes and imagined events.

In an empirical case study, Johnson, O'Connor & James (1997) report a confabulating patient (GS) with impaired source monitoring. Compared to three non-confabulating patients with frontal lobe dysfunction and three non-brain injured ‘normals’, GS tended to underestimate the passage of time and had impaired ability to remember which of two speakers had presented information. He also produced as much detail during confabulatory recall as when accurately recalling an event. They concluded that his confabulation was due to the combination of deficient autobiographical memory retrieval, source monitoring deficits and a propensity toward detailed imaginations. However, temporal source monitoring appears to be intact as his performance was normal for temporal discrimination. This ability to remember which memories are more
recent or distant seems more relevant to confabulation than the ability to estimate how long a task has taken. His normal performance on one of the two recognition tasks raises concern about the reliability of their method.

Johnson's proposal could be tested by studying the relationship between the nature of impaired reality monitoring and confabulation errors. For example, in their case study the dissociation between monitoring temporal and speaker source should be mirrored in GS’s confabulations. He should be more likely to confabulate about who said or did things than to misplace events in temporal context. Johnson et al’s 'combined deficits' account could be tested in a similar way. For example, GS’s confabulations suggest combined deficits of autobiographical memory and internal source monitoring. He recalled few details for autobiographical memories yet his confabulations were as detailed as accurately recalled memories suggesting that internally sourced information substituted for the missing details. Although no empirical measure of internal source monitoring was used, this might have been done using Johnson & Raye's (1981) method on which we would expect GS to show an abnormally enhanced 'generation effect' for imagined versus actual events.

An impaired comparator process could also be due to the use of inappropriate criteria when comparing the memory trace and target specification. Over-reliance on familiarity information has been suggested as a cause of false memories. Evidence of abnormally high levels of false-positive errors and memory judgements based on 'knowing' rather than 'remembering' are suggested indicators of reliance on familiarity information (Parkin 1997, Hanley & Davies 1997). These are likely to reflect the persistence of an impairment that contributed to the confabulations in the patients studied even though they were no longer confabulating at the time of the experiments.
Because Parkin suggests impaired encoding causes reliance on familiarity information, a different explanation is needed for confabulation involving premorbid memories. Familiarity information may be more accessible or influential during retrieval. Recognition memory tests tend to be more resilient to memory impairment (Aggleton & Shaw 1996) and they are said to reflect the use of familiarity information (see Parkin 1997). Familiarity features may be accessed by associative memory processes, that are not thought to be impaired in confabulation. Combined with impulsiveness, that is also said to contribute to confabulation (Stuss et al 1978, Kapur and Coughlan 1980, Moscovitch & Melo 1997) reliance on familiarity could produce confabulations. This could be explored using a Stroop-type measure to assess interference from familiarity. Evidence of interference from a more recent word list while judging whether words were from an earlier list would support this.

*Output Monitoring Processes*

Even if other processes were impaired confabulation would be unlikely if output monitoring processes were intact. Impaired monitoring is suggested to differentiate confabulators from non-confabulators (Moscovitch & Melo 1997). Although neuro-imaging evidence suggests that output monitoring and strategic retrieval are dissociable, there is no current behavioural evidence in support nor any obvious methodology.

Demonstrating the separate existence of an output monitor and a comparator is challenging. In theory if either is intact there would be no errors. This seems unlikely however, as there is evidence that making cognitive errors is normal (e.g. Reason 1979), including during recollection (Burgess & Shallice 1996a). It has been suggested that by its design as an iterative process that is monitored and regulated, output will sometimes be flawed but usually 'sufficient' (Gigerenzer & Goldstein 1996). Hence even if one process is intact there will be errors that can be studied.
Frontal Lobe Dysfunction

Frontal lobe dysfunction (FLD) or ‘dysexecutive syndrome’ (DES) has been implicated in confabulation (e.g. Baddeley & Wilson 1986). DeLuca & Cicerone (1991) found spontaneous confabulation only in patients with ACoAA\(^6\) which causes frontal lobe damage. Unfortunately they did not include measures of FLD and their analysis of lesion locus was not specific. Tests of FLD were included by Moscovitch & Melo (1997) and with more detailed lesion analysis they were able to implicate specifically the ventro-medial frontal area. Understanding the anatomical substrate allows us to extrapolate to function as Moscovitch & Melo did by referring to studies that link the ventro-medial cortices to strategic retrieval processes such as monitoring.

There is support for the suggestion that the severity of FLD determines whether confabulation is ‘spontaneous’ or ‘fantastic’ (e.g. Stuss et al 1978). However, Kapur and Coughlan’s (1980) description of a patient whose fantastic confabulation ceased with improved performance on tests of FLD is questionable as some of the measures may improve with repeated administration (e.g. the Wisconsin Card Sort; Lezak 1995). Fortunately their conclusions were supported by evidence of behavioural improvement. Similar findings in a patient reported by Papagno & Baddeley (1997) although “his executive problems had not completely resolved”. The nature of these was unclear as no concurrent tests were reported. As behavioural signs of DES can occur in the absence of impairment on traditional tests of FLD (Shallice & Burgess 1991) results of these tests alone are not sufficient evidence on which to reach conclusions about the relationship between FLD and confabulation.

Confabulation is not always associated with impairment on standard clinical tests of FLD. One non-confabulating patient with bilateral frontal damage reported by Delbecq-Derouesné, Beauvois, & Shallice (1990) scored in the

\(^6\) Anterior Communicating Artery Aneurysm.
normal range on various FLD tests. It was also reported that Dalla Barba’s (1993a) patient with BE did not show any evidence of frontal pathology. This is doubtful not only because his performance was poor on the ‘verbal fluency’ test which is associated with FLD but also because BE causes widespread subcortical damage causing DES. The significance of verbal fluency performance and confabulation is not clear as it is sometimes impaired (Kopelman et al. 1997; Cunningham et al. 1997) but other times normal (Papagno & Baddeley 1997; Kapur & Coughlan 1980).

The assertion that FLD contributes to confabulation has two weaknesses. FLD is not a unitary deficit but rather encompasses a variety of cognitive functions (Burgess & Shallice 1994) and as mentioned, traditional tests of FLD are not sufficiently reliable indicators of FLD (Shallice and Burgess 1991). Hence as Johnson et al. (1997) conclude “our understanding might be increased by identifying confabulating and non-confabulating patients who resemble one another on some neuropsychological tasks and compare them on additional tasks designed to assess the more specific cognitive factors that have been proposed to underlie confabulation”.

**Personality**

Although theories suggesting that personality determines confabulation (Weinstein 1991) have tended to be dismissed (Moscovitch 1995) recent evidence suggests this may be premature. Loftus (1998) found that “a single act of imagining a known counter-factual event can increase the subjective likelihood that the event would be remembered as having happened in the past”. Only found in one in three people, such ‘imagination inflation’ was associated with higher scores on the Dissociative Experiences Scale. Although an interesting area to explore, it will be necessary to overcome the difficulty of reliably assessing pre-morbid personality retrospectively.

---

7 Binswanger’s encephalopathy.
COMBINATION OF FACTORS

Clearly from this review a single factor is not responsible for confabulation. Consequently, several researchers have suggested that confabulation results from a combination of factors (Cunningham et al 1997, Kopelman et al 1997, Johnson et al 1997).

Burgess & Shallice (1996a) proposed a model combining the factors mentioned above following a study of the normal autobiographical recall process. From the protocols this produced they identified various ‘elements’ (e.g. Memories, Corrections etc.) and then analysing the sequential relationships between them to form four groups of elements. The Editing Processes tended to occur around the Memory elements, providing ‘on line’ editing of the retrieved memories. Preceding editing and memory elements the Descriptors of Recall Requirements, including target specification, describe what is required in recall while the Strategic Mediator Processes, such as ‘task demand analysis’, are involved in problem solving. Although usually rectified, errors were common, and were similar to confabulations, including production of incomplete memories that were claimed to be complete and recall of unrelated but inferentially linked memories. Verification of memories took place throughout the process in various forms including assessing whether the solution fitted the requirements and checking whether recollections were consistent and plausible. The model that they then propose includes the long-term memory storage system, memory control processes (i.e. descriptor, mediator and editor) and other cognitive systems that are involved in problem solving. They argue that confabulation can arise from impairment in any of the three memory control processes.

This is an exploratory study that helps to develop a model. Its limitation is that has not tested a priori hypotheses. For further empirical work clearer operational definitions and measures of the control processes need to be developed as well as methodology to allow them to be manipulated independently.

CONCLUSIONS

Confabulation is most likely due to the combination of memory control and executive processes. In order to study their relative contributions suitable a methodology is needed. This includes improving the reliability of measures such as the Confabulation Battery. Experimental designs are also needed that can explore the combined effects of several factors.

Various ideas about possible future research have been suggested. The Confabulation Battery should be modified and compared with the cue-word method regarding semantic/episodic and narrative/non-narrative memories. Comparing the effects of 'old' versus 'new' cues may help to explore the search initiation versus target specification distinction. The pattern of deficient and intact processes should be mapped using several experimental measures and then compared to the nature of confabulations. The hypothesis that certain informational features (e.g. familiarity) are more influential could be tested using a form of the Stroop effect. We also need to focus on specific frontal functions using specially designed measures. Finally, we should consider how the issue of personality might be included in research. Such empirical work is needed in order to develop treatment methods.
REFERENCES


A Review of the Clinical, Conceptual and Empirical Adequacy of Measures of Awareness Deficit Associated with Frontal Lobe Dysfunction

Introduction

In addition to causing various types of cognitive impairment, brain injury can also affect the capacity to recognise that an impairment exists. Damage in various areas of the brain can cause this deficit for which a variety of terms have been used including anosognosia, lack of insight, denial of illness and unawareness of deficits (see McGlynn & Schacter 1989). A variety of assessment measures are also now available.

Development of adequate measures of unawareness is of both academic and clinical interest. Deficient awareness is known to be associated with unsuccessful vocational rehabilitation (Ben-Yishay & Prigatano 1990). It also contributes to difficulties in psychosocial adjustment (e.g. Oddy et al 1985) which often become chronic (see Oddy 1993). Adequate measures are needed for use in rehabilitation assessment and outcome evaluation. They are also needed for use in research of this phenomenon.

This essay will focus on the form of unawareness associated with frontal lobe damage. Because adequate measures must take into account the clinical presentation of such unawareness as well as the theoretical concepts and theories these will be summarised before reviewing the various available measures.
Clinical Examples of Awareness Deficit

Some clinical examples will illustrate the nature and variety of the phenomena. Although these can be found in the literature (e.g. Prigatano 1991a,b) I will describe some from my own clinical experience.

After suffering a severe TBI\(^1\), in a traffic accident FB’s memory and other cognitive functions were severely impaired. He did not report any deficits although when asked specifically he acknowledged some, such as acquired dysgraphia. He described his writing as “messy” but could not understand why this occurred other than to suggest that his writing ability had never been good. He never expressed concern or distress about the deficits whose severity he grossly underestimated and when they were pointed out he would blandly accept what was said or chuckle and appear bemused. He did not notice when they occurred and did not attempt any adjustment. His expectations of being able to live independently and return to work were completely unrealistic. He attended daily for rehabilitation “for enjoyment” rather than to improve his abilities.

Although DC was aware that he suffered cognitive impairments following a TBI, he did not accept his wife’s suggestion that his behaviour had changed. She reported that he bathed and changed into fresh clothing less frequently but he felt that he did these as often as before and that his standards had not altered. Similarly, he did not agree with her view that he was less affectionate and emotionally demonstrative. JW, who also suffered a severe TBI while riding his bicycle, accepted the possibility that he had become more strict and authoritarian and no longer showed humour when interacting with his young children, but failed to notice when this was happening. He was also unaware that during conversations with friends he rarely showed an interest in their lives’. When pointed out to him he said it was because he could not remember

\(^1\) Traumatic brain injury (e.g. closed head injury).
any personal information about them. However, there was minimal retrograde amnesia, his memory test scores were in the average range and he had an excellent memory for daily experience.

Although JS acknowledged that he had been injured during an assault, he did not accept that he had had a TBI particularly as he did not feel that he experienced any of the sequelae that were described to him or that he observed in the other patients. He failed to notice when he spoke in excessive detail or length and did not detect others’ non-verbal reactions to this. He also failed to notice others’ discomfort when he stood too near to them or spoke with inappropriate familiarity or intimacy. He felt able to return to work well before the therapists did. Similarly, LG felt able to return to nursing after suffering diffuse cerebral anoxia. She was unaware that her comprehension was so impaired that she often failed to understand what people said. She knew that she could not understand the names of foodstuffs but she did not foresee how this might limit her ability to look after two young children.

These examples illustrate various facets of awareness deficit such as failure to recognise or acknowledge the existence of deficits. They often fail to monitor themselves and thus fail to detect when deficits affect their performance. With this there is the failure to compensate or to adjust in some way for the deficits. This can be associated with a lack of emotional acknowledgement in the form of lack of concern and indifference. Deficient awareness may include failure to understand the link between HI, TBI and deficits, in some cases attributing these to ‘natural’ or premorbid factors. The person may also believe that rehabilitation is no longer relevant and that they are capable of resuming responsibilities such as working and parenting. There are also the consequences for social interaction. They may be unaware of emotional changes such as ‘coldness’ or being less affectionate and emotionally expressive. They may be less interested and concerned in others and less able to sense other’s feelings and reactions to their behaviour, such as inappropriate proximity or intimacy.

Another aspect of altered behaviour about which they may be unaware is an alteration in personal standards, such as personal hygiene.

It is in order to account for these phenomena that we need to develop a conceptual framework, models and adequate measures of awareness. Therefore we will next consider the ways in which they have been conceptualised in the literature.

**Conceptualising Frontal Awareness Deficit**

Unawareness deficit can occur in a variety of forms, as McGlynn & Schacter (1989) illustrated in their review such as unawareness of neglect, cortical blindness, hemiplegia and amnesia. Such forms of ‘unawareness of deficit’, associated with posterior damage, are said to differ from impaired self-awareness that occurs with frontal damage (Stuss 1991a,b, Stuss & Benson 1986, Stuss, Picton & Alexander in press). The key question is what constitutes the ‘self’ and self-awareness. Therefore this section will focus on the main issues in conceptualising self-awareness. These have been drawn from the writings of Stuss (op cit.) and Prigatano (1991a, Prigatano & Schacter 1991). Stuss has identified three key components of self and self-awareness; creation of the self-image, monitoring of the self’s condition and awareness of the self.

Several aspects of knowing one’s self have been mentioned in the literature. One is the ability to be aware of one’s own thoughts and emotions. This requires an awareness of internally generated information and the ability to distinguish it from external input (Johnson 1991). It has also been described as the ability to “maintain a sense of subjectivity”, “a private . . unique interpretation of an experience” (Prigatano 1991a). Self-monitoring is involved in both the capacity for ‘self-reflectiveness’ and the spontaneous initiation of such ‘reflection’.

As well as the 'sense of subjectivity', self-awareness is the "the capacity to perceive the 'self' in relatively objective terms" (Prigatano 1991a). Objectivity is defined as "perceiving a situation, object or interaction in a manner quite similar to others' perception". One aspect of this is judging the accuracy or veracity of perception using abstract abilities such as logic and rational reasoning. This assumes that there is a 'true' view of 'reality' which others will perceive. The second concerns the ability to 'know' another person's view. Although not suggested in this literature, this may involve 'theory of mind' (Leslie 1994, Baron-Cohen 1995) as it relates to the ability to form an abstract representation of another person's thoughts, feelings and knowledge and to distinguish this from one's own.

Self-awareness is also said to involve social perception. In order to be aware of the impression one is making, a person must be able to detect and accurately interpret others' reactions. It therefore includes awareness of how one's behaviour affects or may affect others. One's self image, or the 'creation of self', arises in part from perceiving the "position of the self within the social milieu" (Prigatano 1991a). "Knowing who one is often depends on knowing who one is in relation to someone else" (Prigatano 1991a). Hence, self-awareness and social-awareness are suggested to go hand in hand.

The social psychological literature (e.g. Duval & Wicklund 1972, Carver & Scheier 1981, 1982, Sedikides 1992, Salovey 1992) provides ideas and models about self-awareness but limited reference is made to this in the neuropsychological literature (e.g. Stuss & Benson 1986). One suggestion is that self-regulation involves self-awareness in the process of feedback and behavioural regulation. Another notion is that attention can either be focused internally or externally with the switch between foci involving detection of important relevant information (e.g. increase in arousal/anxiety, changes in other's reactions). Self-awareness is also said to involve a set of 'standards' against which current behaviour is compared to determine whether there is a
discrepancy that needs behavioural rectification. This seems relevant to changes in self-standards that can follow a TBI (e.g. poorer hygiene).

Self-image and self-awareness also involve the notion of self-efficacy. This refers to “beliefs in one's capabilities to mobilise the motivation, cognitive resources and course of action needed to meet given situational demands” (e.g. Bandura 1989). Considering that underestimation of limitations is common after frontal damage it is surprising that this is not mentioned in the neuropsychological literature. Underestimation usually leads to unrealistic expectations, such as planning to resume work. The involvement of goals and ‘intentions’ in self-awareness has been mentioned by Prigatano (1991a) who wrote that “the capacity or the state to intend some action . . constitutes the ‘stuff’ out of which self-consciousness or self-awareness emerges”. A person’s ‘self’ is said to be manifest in their goals and therefore in their ability to control their behaviour such that they can be fulfilled.

Schacter & Prigatano (1991) have identified several issues that also need to be taken into account in designing measures. One is that awareness can occur at different levels. Crosson, Barco, Velozo, et al. (1991) proposed three different levels of increasing awareness. ‘Intellectual awareness’ is the knowledge or understanding that a function is impaired. ‘Emergent awareness’ exists when this is coupled with the ability to recognise when a problem is actually occurring. At the highest level ‘anticipatory awareness’ involves anticipating that a problem is likely to occur and hence preparing for this (e.g. using a method to compensate for the impairment). Secondly, awareness may be partial. For example, a person’s actions may indicate that they have implicit awareness even though they do not show explicit awareness of their impairment. Thirdly, awareness can be specific. A person may be aware of some of their deficits while remaining unaware of certain others. Finally, there is the issue of ‘defensive denial’. Prigatano (1991b) has defined denial as a psychological defence mechanism whereas unawareness is an organic deficit.
In summary, awareness deficit associated with frontal dysfunction differs from other forms of unawareness as it concerns impaired awareness of 'self'. The issue of how self-awareness is defined has been discussed by some authors but its conceptualisation is not yet sufficiently clear. Numerous facets have been mentioned and these are summarised in table 1. These need to be taken into account in the design of self-awareness measures.

Table 1
Concepts and Facets Involved in Self-awareness

<table>
<thead>
<tr>
<th>Internal Awareness</th>
<th>Social Awareness</th>
</tr>
</thead>
<tbody>
<tr>
<td>affective state</td>
<td>social standards</td>
</tr>
<tr>
<td>thoughts</td>
<td></td>
</tr>
<tr>
<td>personal standards</td>
<td></td>
</tr>
<tr>
<td>spontaneous self-reflection</td>
<td></td>
</tr>
<tr>
<td>subjectivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>External Awareness</td>
<td></td>
</tr>
<tr>
<td>objectivity</td>
<td></td>
</tr>
<tr>
<td>other's mood</td>
<td></td>
</tr>
<tr>
<td>other's thoughts</td>
<td></td>
</tr>
<tr>
<td>other's behaviour</td>
<td></td>
</tr>
</tbody>
</table>

Theories of Frontal Awareness Deficit

McGlynn & Schacter (1989) provide a good review and summaries of various theories. However, most relate to awareness deficits associated with posterior damage, such as hemiplegia and visuo-perceptual impairment, rather than unawareness associated with anterior damage.

There are several proposals based on disconnection, usually between the contents of a system (e.g. sensation, vision, proprioception) and the mechanism that enables conscious access to its contents. Disconnection of the body image representation from awareness is one account of unawareness of hemiplegia (Gerstmann 1942; cited in McGlynn & Schacter 1989), and disconnection of the 'speech area' from the region allowing 'introspection' about speech.
performance was proposed for unawareness of dysphasia (Geshwind 1965; op cit.). A different type of disconnection was proposed by Halligan & Marshall (1998) to explain visual neglect of the left half of space. They proposed a disconnection of two visual processing mechanisms (focal and global) causing a loss of awareness of the deficiency in global processing (i.e. the failure to maintain the ‘panorama’ that includes both left and right visual fields).

Disconnection is also implicated in some accounts of frontal awareness deficit (Schacter 1989, 1991; McGlynn & Schacter 1989; Stuss 1991a,b; Stuss, Picton & Alexander in press). Schacter suggests that awareness occurs as a consequence of input to the ‘conscious awareness system’ (CAS) from various modules (e.g. linguistic and perceptual). Awareness deficit results when a module is disconnected from the CAS. Disconnection of frontal areas would remove “awareness of more complex deficits, such as difficulties in problem-solving, retrieving and integrating information as well as social, behavioural and personality changes” (McGlynn & Schacter 1989).

A disconnection between frontal and parietal processes was proposed by Stuss to explain impaired self-awareness. Parietal functions focus on external information (i.e. sensori-perceptual) while frontally the focus is on internally generated information. In his model the highest level of frontal functioning is self awareness. This involves integration of abstract mental representations of the past (i.e. personal memories), the present (i.e. the “self’s condition”, including emotion, etc.), the future (e.g. intentions, goals, expectations) and knowledge of the world (i.e. semantic memories). Normally there is a balance of focus between internal and external factors through frontal inhibition of parietal activity. Disconnection of these regions causes an imbalance resulting in “an excessive approach to the environment, leading to . . . an excessive distance from the intrapsychic processes necessary for insight, foresight and abstraction”. This is similar to ‘utilization’ behaviour (Shallice, Burgess, Schon & Baxter 1998) and awareness deficit.
The integration of information from several sources is another explanation proposed for awareness. Unawareness of hemiplegia was suggested to arise from failure to integrate sensory information (Critchley 1949; cited in McGlynn & Schacter 1989). As mentioned above, Stuss suggested that failure to integrate information about past, present and future could result in self-awareness deficit. Prigatano has argued that self-awareness deficit results from impaired integration of extra-personal and intra-personal information. He argued that a frontal injury damages heteromodal cortex which integrates extra-personal information, such as perceptions of other’s behaviour, with intra-personal information, such as the person’s own affective reaction. He also suggests that impaired integration can cause unawareness by interfering with the integration of thoughts and feelings, self-perception, social awareness and impairment of intention and drive.

Impaired monitoring is another account of unawareness (Johnson 1991, Johnson, Hastroudi, Lindsay 1993, Johnson, O’Connor & Cantor 1997). In the Multiple-Entry Modular (MEM) model, Johnson has distinguished posterior ‘perceptual’ and anterior ‘reflective’ systems. The reflective system contains ‘supervisory’ processes (e.g. ‘initiating’) and ‘executive’ processes (e.g. ‘shifting’). The interaction between these is said to result in our ‘concept of self’ and ‘sense of self’ as well as the abilities of being aware of our cognitive processes and ‘reflecting on reflection’. They are also crucial to ‘reality testing’ which monitors the source of information (e.g. internal/external) as well as the consistency and plausibility of information. Because supervisory and executive processes monitor one another, impairment of either or a disconnection of the two would disrupt self-awareness. Although the nature of this is not specified, we can infer that ‘reality testing’ will be defective and therefore any discrepancy between current and remote memories will not be detected. For example, any discrepancy between current and past performance would not be detected thus resulting in unawareness of change.
In summary, measures of awareness need to be capable of exploring the proposals of disconnection, impaired integration and defective monitoring. Measure are needed that can distinguish different facets of awareness. These include awareness of personal state (e.g. affect), current and past performance and others' social behaviour (e.g. non-verbal behaviours). Primarily the measures must provide a valid and reliable indication of the individual’s specific form of awareness deficit so that empirical measures can be used to test the specific hypothesis relating to that form of unawareness.

Adequacy of Awareness Deficit Measures For Research and Clinical Use

There are a variety of measures that have been used to assess unawareness deficit. These are displayed in table 2 which provides basic information about each including the type of measure and number of items.

Unawareness is usually operationalised by measuring the discrepancy between the patient’s self report and some standard. This is usually done by comparing self-rated performance with the ratings of a carer\(^2\) (e.g. Prigatano et al 1986, Godfrey et al 1993) or a clinician (e.g. Fordyce & Roueche 1986, Gasquoine 1992, Zhou et al 1996). The weakness of this approach is that the carer’s or clinician’s ratings may not be accurate. Carers may ‘deny’ the severity of a relative’s problems (Roman 1974; cited in Sherer et al. 1998b). Alternatively, their personality characteristics may lead to an inflation of severity (McKinlay & Brooks 1984). The clinician’s emotional reaction to the patient can also affect their rating of severity (Heilbronner, Roueche, Everson & Epler 1989; cited in Sherer et al. 1998b). Although the correlation between carers’ and clinicians’ ratings are greater than those between carer and patient or clinician and patient, they are still far from perfect (Fordyce & Roueche 1986). It is

---

\(^2\) The term carer will be used here to indicate either relative or staff who are involved with support or supervision of the person with a BI.
<table>
<thead>
<tr>
<th>Measure</th>
<th>Author</th>
<th>Items</th>
<th>Rating</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Everyday Memory Questionnaire</td>
<td>(EMQ) Sunderland et al (1983)</td>
<td>28</td>
<td>S</td>
<td>C</td>
</tr>
<tr>
<td>Patient Competency Rating Scale</td>
<td>(PCRS) Fordyce &amp; Roueche (1986)</td>
<td>30</td>
<td>S</td>
<td>C / S</td>
</tr>
<tr>
<td>Change Questionnaire</td>
<td>(CQ) Hendryx (1989)</td>
<td>15</td>
<td>S</td>
<td>C / S</td>
</tr>
<tr>
<td>Awareness Interview</td>
<td>(AI) Anderson &amp; Tranel (1989)</td>
<td>8</td>
<td>QA</td>
<td>S / T</td>
</tr>
<tr>
<td>Employment Rehab Centre scale</td>
<td>(ERC) Herbert &amp; Powell (1989)</td>
<td>12</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>San Diego Questionnaire</td>
<td>(SDQ) Allen &amp; Ruff (1990)</td>
<td>29</td>
<td>S</td>
<td>T</td>
</tr>
<tr>
<td>Cognitive Rating Scale</td>
<td>(CogRS) Gasquione (1992)</td>
<td>10</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Anosognosia Questionnaire</td>
<td>(AnQ) Starkstein et al (1992)</td>
<td>6</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Head Injury Behaviour Scale</td>
<td>(HIBS) Godfrey et al (1993)</td>
<td>27</td>
<td>S</td>
<td>C</td>
</tr>
<tr>
<td>Self-Awareness Questionnaire</td>
<td>(SAQ) Gasquione &amp; Gibbons (1994)</td>
<td>27</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Multiple Ability Self-report Questionnaire</td>
<td>(MASQ) Seidenberg et al. (1994)</td>
<td>38</td>
<td>S</td>
<td>C / S</td>
</tr>
<tr>
<td>Awareness Questionnaire</td>
<td>(AQ) Sherer et al (1995)</td>
<td>18</td>
<td>S</td>
<td>C / S</td>
</tr>
<tr>
<td>‘ad hoc’ Questionnaire</td>
<td>(ad hoc) Hütter &amp; Gilsbach (1995)</td>
<td>8</td>
<td>S</td>
<td>C</td>
</tr>
<tr>
<td>Competency Rating Scale</td>
<td>(ComRS) Zhou et al (1996)</td>
<td>47</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Knowledge of Acquired Brain Injury</td>
<td>(KABI) Zhou et al (1996)</td>
<td>N/A</td>
<td>N/A</td>
<td>facts</td>
</tr>
<tr>
<td>Self-Awareness of Deficits Interview</td>
<td>(SADI) Fleming et al (1996)</td>
<td>3</td>
<td>QA</td>
<td>N / A</td>
</tr>
<tr>
<td>Impaired Self Awareness</td>
<td>(ISA) Prigatano &amp; Klonff (1998)</td>
<td>10</td>
<td>C</td>
<td>N / A</td>
</tr>
<tr>
<td>Defensive Denial</td>
<td>(DD) Prigatano &amp; Klonff (1998)</td>
<td>10</td>
<td>C</td>
<td>N / A</td>
</tr>
<tr>
<td>Structured Awareness Interview</td>
<td>(SAI) Giacino &amp; Cicerone (1998)</td>
<td>N / A</td>
<td>QA</td>
<td>C</td>
</tr>
</tbody>
</table>

1 Rating by Self (S) or Clinician (C) or answers to questions (QA)
2 Comparison criterion; carer (C), staff/clinician (S), tests (T), performance/observation (P / O)
clearly difficult to obtain reliable ratings of a patient’s abilities for use as criteria against which to measure the degree of awareness deficit.

Other forms of clinician ratings have also been used. In the Awareness Interview ratings are based on answers given during a structured interview (Anderson & Tranel 1989). One advantage of this method is that it reduces the possibility that the patient may not understand the questions. The wording on written questionnaires can influence the patient’s rating. Gasquoine (1992) compared two questionnaires, one worded using clinical terms for impairments and the other using descriptions of everyday activities to illustrate the impairments (e.g. “concentration” versus “the ability to stick to a task such as reading a newspaper”). The level of agreement between the two questionnaires was 73% for specific questions (e.g. “Has your memory been affected by the accident?”) and only 54% for general questions (e.g. “How is your memory now?”). The potential disadvantage of Anderson & Tranel’s method is inaccuracy by the clinician. However, they reported good inter-rater reliability (r=0.92).

The wording of questions has been a neglected issue in this area of questionnaire development. There is considerable variety amongst the methods. Some ask whether the activity/ability is ‘troublesome’ (Hüttter & Gilsbach 1995) or is a ‘problem’ (Allen & Ruff 1990, Godfrey et al 1993) or alternatively how often they find it problematic (Seidenberg et al. 1994, Zhou et al. 1996). Others ask whether their ability has ‘changed’ (Hendryx 1989) or is ‘better/worse’ (Sherer et al 1998). The degree of ‘ease’ (Fordyce & Roueche 1986), ‘difficulty’ (Hart et al 1998) or level of ability (Gasquoine 1992) is asked in others. A visual analogue scale has also been used to indicate level of performance (Herbert & Powell 1989).
Clinical experience indicates that when asking patients about the sequelae of a brain injury the wording of the question has a strong influence on their response. Some wording is more negative (e.g. "what problems do you have?") or likely to imply inadequacy (e.g. "what do you have difficulty doing?"). Other wording can be less challenging and more likely to elicit reports such as asking about "changes" or "what has not fully recovered yet?". The specificity of the question is also important. Open, general questions are likely to elicit fewer responses indicating any awareness than more specific questions. For example, Gasquoine (1992), like others, found that broad questions (e.g. "What problems have you had since the accident?") elicited reports of problems in only about half of the patients. However, gradually increasing question specificity may be a useful method in itself, as the resultant increase in detail can reflect levels of awareness. Research is needed to examine both the effect of question wording and the relationship between the information obtained and severity of unawareness.

Another method of clinician rating is global rating of awareness (Fleming et al 1996, Sherer et al. 1998c). Although this is likely to be reliable due to its generality, it provides virtually no detailed information about the nature of any awareness deficit. Hence this method would not be useful for exploring particular forms of unawareness (e.g. awareness of internal state or others’ reactions).

Patient’s self-ratings have also been compared with the results of tests (Anderson & Tranel 1989, Allen & Ruff 1990). This has the advantage of objective criteria. Although this practice is more common in studies seeking to validate an awareness measure (e.g. O’Donnell et al. 1983), Allen and Ruff used it in their comparison of mild and severe head injured patients. However, there can sometimes be a limited degree of correspondence between test performance and everyday functioning. (Heaton & Pendleton 1988). Another complication is the question of what constitutes an ‘abnormal’ correlation.
between self-rating and test performance in view of the finding that the non-brain group in Allen & Ruff's study overestimated their performance on some tests (e.g. logical thinking).

The strength of comparing self-rating of task performance and behavioural observation of the task is that the experience on which the patient is basing their judgements is known. This approach was used by Hart, Giovanetti, Montgomery & Schwartz (1998) to study awareness of errors that occur in an everyday task (e.g. making a sandwich). However, they did not find that TBI patients overestimated their performance significantly more than the control group. One potential draw-back of this method is that the close juxtaposition of performance and self-assessment may produce a different level of awareness than occurs when the rating relates to an open ended period of time, as most do. Shallice & Burgess (1991) have argued that the structure of assessment methods may counteract the executive deficit in question. Consequently, evidence of frontal dysfunction may be minimised or masked.

Another advantage of comparing task performance with self-rating of performance is that it allows a focus on a particular facet of awareness. Azouvi, Marchal, Samuel et al. (1996) used this method to explore awareness of unilateral neglect. In addition to comparing self-rating with performance on tests of neglect (e.g. cancellation tests), they also used a clinician rating scale based on behavioural observations in daily activities (e.g. eating and dressing). An even more focused measure was used by Carpenter, Berti, Oxbury, et al. (1995) who asked patients about their ability to move their arms during sodium amytal induced paralysis. The question is whether it would be possible to design such a task to elicit less tangible forms of unawareness (e.g. unawareness of others' emotional reactions).
There is a wide range of number of items in the various measures (see table 2). There is also considerable variation in the areas covered by these measures (see table 3). Most, but not all, cover cognitive effects of brain injury. However, there is also considerable variation in which aspects of cognition are included. Many apparently do not include memory (e.g. Hütter & Gilsbach 1995) or it is unclear due to lack of detail (e.g. Godfrey et al. 1993). Nevertheless, there is little consistency in the item composition. Measures that do not include memory, however, would not be suitable for TBI patients or even patients with frontal damage (e.g. ACoAA/SAH\(^3\)) which can cause memory impairment (Baddeley & Wilson 1988).

Item heterogeneity is not necessarily a problem, particularly if the measure was designed to assess a specific deficit, such as memory (Sunderland et al. 1983) or neglect (Azouvi et al. 1996), or for a particular group of patients (e.g. stroke; Starkstein et al 1992). However, item heterogeneity is needed to address the issue of ‘specificity’ of awareness (Prigatano & Schacter 1991). Without covering various areas it would not be possible to detect those in which the person’s awareness was intact and those where it was deficient. Although more comprehensive measures are likely to be needed for clinical purposes, focused measures will often be needed for research on specific patient groups or deficits.

Some measures have grouped their items (e.g. Hendryx 1989, Herbert & Powell 1989, Seidenberg et al. 1994), but few have actually tested the validity of their categories. Although in the original design of the Awareness Questionnaire (AQ) Sherer et al (1998a) proposed four factors (Physical, Behavioural, Cognitive, Functioning in Community Activities), only three factors were revealed by an item factor analysis (Motor/Sensory, Cognitive, Behavioural/Emotional). In view of this discrepancy, item factor analysis is recommended with all self-rating measures to test the validity of the proposed

---

\(^3\) Anterior communicating artery aneurysm / subarachnoid haemorrhage.
<table>
<thead>
<tr>
<th>Area covered in measures of awareness</th>
<th>Sen-Mt</th>
<th>Pcpt</th>
<th>Cog</th>
<th>Exec</th>
<th>P&amp;S</th>
<th>Beh</th>
<th>Soc</th>
<th>Mood</th>
<th>ADL</th>
<th>Other Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuropsych’s Impairment Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Little detail available)</td>
</tr>
<tr>
<td>Everyday Memory Questionnaire</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Main focus is memory</td>
</tr>
<tr>
<td>Patient Competency Rating Scale</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>Work, driving</td>
</tr>
<tr>
<td>Change Questionnaire</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Physical</td>
</tr>
<tr>
<td>Awareness Interview</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Physical</td>
</tr>
<tr>
<td>Employment Rehab Centre scale</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>San Diego Questionnaire</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive Rating Scale</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anosognosia Questionnaire</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Communication, BI¹</td>
</tr>
<tr>
<td>Head Injury Behaviour Scale</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>(Little detail available)</td>
</tr>
<tr>
<td>Self-Awareness Questionnaire</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>BI</td>
</tr>
<tr>
<td>Multiple Ability Self-report Q’naire</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Awareness Questionnaire</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘ad hoc’ Questionnaire</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Competency Rating Scale</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge of ABI²</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>Focus on neglect.</td>
</tr>
<tr>
<td>Catherine Berego Scale</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>Implications of deficits &amp; realism of goals</td>
</tr>
<tr>
<td>Self-Awareness of Deficits Interview</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impaired Self Awareness (ISA)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>{Affect and behaviour</td>
</tr>
<tr>
<td>Defensive Denial (DD)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>{ during interview</td>
</tr>
<tr>
<td>Structured Awareness Interview</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>Asked implications</td>
</tr>
<tr>
<td>Multi-Level Action Test</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ BI- awareness of brain injury/condition/reason for hospitalisation. ² - Acquired Brain Injury

Sen-Mt - sensori-motor; Pcpt - perception; Cog - cognition; Exec - executive; P&S - Problem solving & reasoning; Behv - behaviour; Soc - social/interpersonal; Mood - affect/emotion; ADL - activities of daily living (personal or domestic);
categories. The solution may not be as straightforward as this, however. The factor analysis done by Allen & Ruff (1990) on the San Diego Questionnaire (SDQ) revealed six factors (Sensori-motor, Attention, Language, Learning & Memory, Arithmetic and Logical Thinking). Considering that the AQ and SDQ have a similar number of items (29 versus 27) it is unclear whether they differ in item content.

Although it is good psychometric practice to assess a measure's reliability and validity, this has not always been done (e.g. Hendryx 1989, Gasquoine 1992, Starkstein et al. 1992, Hütter & Gilsbach 1995, Zhou et al. 1996, Giacino & Cicerone 1998). Inter-rater reliability is not always relevant. For example, split-half reliability would be more suitable for the Self-Awareness Questionnaire (Gasquoine & Gibbons 1994) which simply requires judging the patient's response to 'yes/no' questions. Where it is reported, reliability is good (e.g. Awareness Interview \( r = 0.92 \), SADI \( r = 0.82 \), PCRS \( r > 0.90 \)). Similarly the validity has not been assessed for all measures. Some authors have tested concurrent validity using neuropsychological tests (Allen & Ruff 1990) or other behavioural measures (e.g. Social Behaviour Assessment; Godfrey et al. 1993), but many have not assessed validity.

In terms of their capacity to measure the facets self-awareness summarised above (see table 1) most of the measures are not adequate. They tend to focus on awareness of sensori-motor and cognitive deficits. Some do address deficits of reasoning and executive functions (e.g. PCRS, MLAT, SDQ, CogRS, SAQ, AI, AQ). Only some address factors such as social and behavioural symptoms (e.g. PCRS, ERC, AQ) which are involved in self-awareness. However, the range of such items can be very limited (e.g. four items in the AQ). The PCRS contains the most social and behavioural items (approximately ten). Awareness of external information such as others' mood, behaviour and attitude, which overlaps with social awareness, is rarely included in these measures. Even the
PCRS which is the exception, has only one such item (i.e. “recognise when something I say or do has upset someone else”).

Several facets of self-awareness are lacking from all measures. These include awareness of changes in social standards (e.g. “Have people who know you well commented on changes in your personal standards?”) awareness of personal feelings and thoughts (e.g. “Is it as easy as it was before the accident for you to notice changes in your emotions?”; “Can you think over your thoughts and views about things as easily now?”). Assessment of inflated self-efficacy and unrealistic goals is also addressed in only a few of the scales (e.g. SAID, AI).

Few of the measures address the issue of levels of awareness (Crosson et al. 1989). One notable exception is the study of Zhou et al (1996) which included a measure of ‘intellectual awareness’ (the Knowledge of Acquired Brain Injury measure) and ‘emergent awareness’ (the Competency Rating Scale). Interestingly, a dissociation between these measures was found in several patients which supports Crosson et al’s proposal. However, as emergent awareness concerns ‘on-line’ awareness of performance, the CRS was limited in being a retrospective measure of emergent awareness. The study of Hart et al. (1998) offers a more adequate methodology for emergent awareness. Their measure of detection of errors did differentiate the control group from the TBI group who showed lower detection and correction rates.

The only measures designed to distinguish psychological ‘denial’ from organic unawareness are the ISA and DD (Prigatano & Klonoff 1998). Although measures show promise, further development is needed to achieve acceptable psychometric characteristics. The ISA was less able to reliably distinguish high/low impaired self-awareness and the inter-rater agreement for ranking of severity of unawareness was poor.
Conclusions

There is considerable variation in the measures of awareness deficit in terms of their focus, number and range of items and evidence of psychometric soundness. Most are not appropriate for measuring self-awareness because they lack the item content to address the facets of impaired self-awareness that are presented clinically and that form the conceptualisation of awareness. Few measures assess awareness of internal and external information. Development of these might be aided by reference to existing measures of self-awareness in other areas of the literature (e.g. Fenigstein, Scheier & Buss 1975, Burnkrant & Page 1984, Piliavin & Charng 1988). There is also a need for further development of a methodology that combines observational information with self-rating measures. In particular these need to focus on social and behavioural facets of self-awareness.

The adequacy of measures for empirical study is limited primarily by their lack of self-awareness assessment. Some of the available measures have good psychometric characteristics (e.g. the PCRS) but this does not compensate for the need to be able to measure specific facets of self-awareness (e.g. awareness of thoughts and feelings).

Finally, the issue of how to conceptualise and operationalise self-awareness needs further attention. The neuropsychological literature has made scarce use of literature in other areas that would help to develop our concept of impaired self-awareness (e.g. Knapp & Deluty 1987, Morin 1993, Hoyle & Sowards 1993). Due to the limitations of operationalising awareness as the discrepancy between self-ratings and those of others, we should also seek an alternative method.
REFERENCES


An Investigation of the Role of Executive Functions in Cognitive Estimation Test (CET) Performance: implications of results from a task of inter-city distance estimation

Mr. Drew Alcott

Submitted for the Doctorate of Psychology (PsychD) in Clinical Psychology Conversion Programme (1999)

Objective: The main objective of the study was to test the view that the Cognitive Estimates Test (CET) reflects the integrity of the frontal, Supervisory System. The Semantic Memory Access impairment account was also tested. A further aim was to increase our understanding of the psychological meaning of performance on estimation tests.

Method & Design: Patients with a brain injury (BI) affecting frontal lobe functioning were compared with non-brain injured participants. In addition to the CET, participants completed a mileage estimation task (the Milest Test) designed to allow provision of a single set of supplementary information to test the prediction that dysexecutive behaviour will interfere with brain injured patients ability to use the information to improve their estimation accuracy. In addition to total CET score the Milest measures included total proportional error, perseverations, bizarre estimates and the squared correlation, slope and intercept from a linear regression of actual against estimated mileage.

Results: Although the BI group was less accurate, contrary to prediction they improved to a similar degree when supplementary information was available. Their level of bizarre responding was also significantly greater but was not reduced when supplementary information available. CET performance did correlate with the main Milest measures. Bizarre responding correlated with the CET in the BI group and it identified more patients whose performance was poor.

Conclusions: There is support for the proposal that estimation performance reflects Supervisory functions, but those of response regulation to a greater extent than those of strategy planning. A theoretical account of estimation is offered that encompasses the results.
An Investigation of the Role of Executive Functions in Cognitive Estimation Test (CET) Performance:
implications of results from a task of inter-city distance estimation

Drew Alcott

CONTENTS

INTRODUCTION
The Cognitive Estimation Test 104
Subsequent Development of the CET 105
Theoretical Account of the CET 107
Sensitivity of the CET to Frontal Lobe Function 109
The Frontal Lobes and Estimation: other evidence 113
Summary 119
Testing the Explanations of CET Performance 120
Aims 122

METHOD
Participants 123
Measures 125
Procedure 125
Design 129
Specific Predictions & Statistical Analysis 129

RESULTS
Effect of Brain Injury on Accuracy of Estimation 132
Ability to Benefit from Supplementary Information 133
Effect of Brain Injury on Dysexecutive Measures 138
Effect of Supplementary Information on Dysexecutive Measures 139
Relationships Between Measures 143

DISCUSSION
Test of the Predictions 146
Critique of the Study: Complications and Limitations
Implications for the Supervisory Model Account of CET Performance
Developing a Theoretical Model

SUMMARY 164

REFERENCES 165

APPENDICES: CET, Milest Test 169
INTRODUCTION

Research into the cognitive functions subserved by the frontal lobes of the brain has continued over the past several decades (e.g. Rylander 1939, Halstead 1940, Goldstein 1944, Reitan 1959, Teuber 1964, Shallice & Evans 1978; see Walsh 1978 for discussion). This has been a process of defining and clarifying the 'executive' functions involved (e.g. see Roberts, Robbins, Weiskrantz 1998). Some of the experimental measures used in such studies have been used clinically to assess frontal lobe damage. Examples of this include the Colour-Form Sorting Test (Weigl 1941), the Category Test (Halstead 1947), Wisconsin Card Sorting Test (Milner 1963) and the Cognitive Estimation Test (Shallice & Evans 1978). The impairments inferred from the results of such tests relate to the cognitive functions said to be involved in the task. This study focuses on the Cognitive Estimation Test (CET) and its theoretical explanation.

The Cognitive Estimation Test

The Cognitive Estimation Test (CET) was designed over 20 years ago by Shallice & Evans (1978) to examine the effect of frontal lobe damage on estimation. After observing grossly defective estimation performance in a patient with severe frontal lobe damage they wanted to explore whether estimation performance was selectively sensitive to frontal lobe damage and whether it was associated with intellectual ability. Their study compared the performance of non-brain injured subjects with patients with different lesions (i.e. "a unilateral focal cortical lesion confirmed to no more than two lobes").

Their task consisted of 15 questions for which the subjects were asked to estimate by "making the best guess" they could to answer each question (e.g. How many slices in a sliced loaf?, What is the length of a pound note?). Questions were designed so as not to require any specialist knowledge and "such that an appropriate plan for answering them was not immediately apparent". They also used tests of arithmetical ability (WAIS Arithmetic
subscale) and general intelligence (Raven’s Matrices) to assess the contribution of these factors to estimation performance.

Data was collected from 96 patients, 45 of whom had damage including the frontal lobes. Patients with frontal injuries were compared to patients with posterior damage and with non-brain injured patients. The control group’s answers were used to determine the criteria for the “extremeness” of the patients’ estimates. Estimates greater than or less than any made in the control group were rated as “very extreme”, those greater or less than all but one of the control group’s were rated “extreme”, and those greater or less than all but two of the control group’s were rated “quite extreme”. These were later said to relate to the 8th, 16th and 25th centiles (Shallice & Burgess 1991a). The groups were compared in relation to the number of extreme responses to explore the relationship between lesion location and estimation performance using analysis of covariance to correct for age, intelligence and arithmetical ability.

They found that patients with frontal damage made significantly more “bizarre” (i.e. “extreme”) responses than patients with posterior lesions. This effect was not dependent upon the overall extent of the lesion or its lateralisation. They also concluded that estimation performance was not determined by either intelligence or arithmetical ability as the main effect of bizarreness was still significant when these factors were used as the covariates.

Subsequent Development of the CET

The CET has subsequently been used extensively in clinical practice and in research. It has also been modified and its psychometric properties have been studied. Perhaps the most commonly used version of the CET is that of Shoqirat, Mayes, MacDonald, Meudell & Pickering (1990). They selected the nine items that were most sensitive to frontal lobe lesions and added the population estimation question from the Information subtest of the Wechsler

Scales (WAIS) to make a total of 10 items. Although their study concerned amnesic patients, their pathologies involved the frontal lobes (i.e. encephalitis, ruptured ACoAA\(^1\) and Korsakoff's syndrome). The patients were significantly worse on the CET than the control group.

A modified and Americanised version of the CET has been designed by Axelrod & Millis (1994) that addresses some of the CET's weaknesses. The bizarreness of certain items can be difficult to determine. For example, the question, "How many camels are there in Holland?", requires the examiner to decide whether an answer constitutes a "large number". It can also be difficult to judge answers to the questions "What is the largest object normally found in a house?" and "What is the best paid job in Britain?" because the scoring guidelines are ambiguous. Neither items specify what answers score a 0 (i.e. 'normal') so this must be decided on the basis of whether the item is larger than a "piano, settee, or sideboard" but not larger than a carpet or whether the person is paid more than "a professional, to and including the PM". Axelrod & Millis developed items that "required numerical responses and were anticipated to have a normal response distribution". Following a normative study the criteria for each item was set in relation to the mean and standard deviation of the sample's responses. Hence for each response a deviation score can be determined and these can be summed to find the total deviation score and to establish the percentile ranking of the person's performance. In a further study Axelrod & Millis compared brain injured patients (severe TBI\(^2\)) with non-brain injured subjects (medical outpatients) using their modified CET and deviation norms. The results for the control group were consistent with those found in the initial study and the TBI patients' CET deviation scores were significantly higher than the controls. This approach is commendable but the test has limited potential for use outside the USA. Several items require culture-specific knowledge (e.g. "How tall is the Empire State building?" and

\(^1\) Anterior communicating artery aneurysm.
\(^2\) Traumatic brain injury.

"""What is the average temperature in Anchorage, Alaska on Christmas Day?""

**Theoretical Account of the CET**

Shallice & Evans (1978) refer to the theoretical position of Luria (1966 op. cit.) to account for the underlying psychological process or mechanism involved in CET performance. The frontal lobes are considered to be the site of cognitive planning, selection and regulation. In their view, performance on the CET relies upon the integrity of such high level, anterior systems. For example, the more ‘routine cognitive skills’ of "mechanical arithmetic calculations" (sited in posterior regions) would be used under the direction of anterior systems which would select an appropriate solution plan and check that the answer was not bizarre.

In a later paper, Shallice & Burgess (1991a) expand upon the argument that abnormal CET performance is due to impaired strategy selection. They proposed that the Supervisory System is involved when circumstances require a response that is not well learned, or which contains novel sequences or actions (Norman & Shallice 1986; op cit.). Regarding the CET they propose that although "no rote knowledge or routine method seems available for obtaining the answer . . . a strategy fairly easily comes to mind that when adopted, quickly provides a reasonable answer". Therefore, they are suggesting that the Supervisory System is involved in formulation of a plan or strategy to answer the CET questions. As a consequence, patients with frontal damage score more highly on the CET (i.e. produce a greater number of bizarre responses) because they are "impaired in devising an appropriate strategy to deal with a problem situation where no routine procedure is available".

The Supervisory process is also said to be involved in evaluating and modifying the outcome of the response plan (Shallice & Burgess 1991a, p. 135). The related explanation of impaired CET performance in frontal patients is that they
“produce bizarre responses which they do not disconfirm” (Burgess & Shallice 1996). In another study which measured bizarre responses, Burgess & Shallice (1996) found that the group of patients with frontal damage produced more bizarre responses (i.e. responses which did not appear to conform to a rule) than those with posterior lesions. One proposed explanation for the findings of Shallice & Evans (1978) and Burgess & Shallice (1996) is that patients with frontal damage “think of strange potential responses . . . and either fail to check their plausibility or are unable (or unwilling) to check their choices”.

A similar, but slightly different account has been suggested by Taylor & O’Carroll (1995). They found that patients with probable mammillary body damage, due to Korsakoff’s syndrome, were significantly worse on the CET than the control group. However, the patients with probable frontal damage (i.e. ACoAA and TBI) were not worse than the normal control group. The mammillary bodies are implicated in memory functions and indeed Korsakoff’s patients tend to have memory impairment with confabulation. Although traditionally these patients demonstrate relatively intact semantic memory, Taylor & O’Carroll refer to evidence that suggests this aspect may in fact be impaired in these patients (Parkin & Leng 1993; op cit.). They argue that the CET requires retrieval of facts/knowledge from semantic memory as well as ‘error checking’. They suggest that the Korsakoff patients’ impaired CET performance may be due to a combination of impaired semantic memory and poor response monitoring associated with a tendency to confabulate.

No significant correlation has been found between CET performance and tests of memory function. In a study of the enhancement of word completion by priming, Mayes and Gooding (1989) found no correlation between memory performance and CET score amongst a group of amnesic patients. Mayes, Meudell & MacDonald (1991) did not find an association between CET performance and two tests of recognition and spatial memory ($r \approx 0.15$) in a study involving amnesic patients. It should be noted that these are tests of new
learning and therefore do not reflect upon semantic memory which is more relevant to the CET (e.g. knowledge about horses, anatomy, Holland, etc.). There is no study exploring the relationship between CET performance and aspects of semantic memory such as retrieval. However, the hypothesis is indirectly supported by more recent studies implicating the frontal lobes in the strategic aspect of retrieval. Evidence from studies of functional imaging and neuropsychological cases suggest that the right frontal region is involved in retrieval (Dolan et al 1997 and Parkin et al. 1997).

Hence, there are a variety of theoretical explanations for impaired CET performance. Two relate to Luria's model of the frontal lobes. Impaired planning and selection of a strategy is one suggestion. The other is impaired regulation of the response (i.e. monitoring and evaluation). Memory impairment (e.g. semantic memory access/retrieval) has also been suggested as a contributory factor. However, there is not been any research published that has aimed to test these. This therefore is the aim of the current study. Before proceeding to describe this however, other relevant literature will be reviewed concerning the CET and estimation tasks.

**Sensitivity of the CET to Frontal Cognitive Functions**

The American CET and its British original have the weakness of considerable item heterogeneity. Estimations cover the domains of height, speed, weight, length, age, pay, population and temperature. Hence it requires knowledge from a variety of domains such as the dimensions of buildings, bottled milk, household objects and the human body. This may affect its psychometric qualities and weaken its clinical power. Axelrod & Millis (1994) found that the degree of correlation between individual items and the total score ranged from 0.22 to 0.57 (p<.001) indicating item heterogeneity. It is not surprising therefore that O’Carroll, Egan & MacKenzie (1994) found poor internal reliability (e.g. split-half reliability; Cronbach’s alpha = 0.40). In addition, their principle component analysis resulted in a five-factor solution for only 10 items.
and only four of the items loaded substantially on the main factor. However, item heterogeneity may be important to the theoretical account of CET performance. The task requirement of "selecting an appropriate plan for answering" the questions may depend upon the fact that a different plan is needed for each item. Reducing item heterogeneity might reduce the task demands on planning and so affect the theoretical implications of the results.

Item heterogeneity may also affect the CET's reliability. A review of studies reporting CET data for 'normal' subjects indicates considerable variation in both the average score and the cut-off score that would be used to indicate 'abnormal' performance (see Table 1). However, the studies also differ in terms of sample size and the form of the CET that was used. Therefore, it would be important to control these variables in further research in order to clarify the effect of item heterogeneity. When the most common form of the CET (Shoqirat et al. 1990) was used with a sample as larger as that collect by

<table>
<thead>
<tr>
<th>Study</th>
<th>mean</th>
<th>s.d.</th>
<th>n</th>
<th>sample source</th>
<th>cut-off¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shallice &amp; Evans &amp; Burgess (1991)*</td>
<td>3.0</td>
<td>2.6</td>
<td>24</td>
<td>no details</td>
<td>8.2</td>
</tr>
<tr>
<td>Leng &amp; Parkin (1988)*</td>
<td>2.9</td>
<td>3.8</td>
<td>7</td>
<td>patients with extra-cerebral lesions; mean age 45 yrs.</td>
<td>10.5</td>
</tr>
<tr>
<td>Kopelman (1988)*</td>
<td>4.5</td>
<td>2.5</td>
<td>16</td>
<td>age 53 yrs. ± 4 yrs.; mean FSIQ 110</td>
<td>9.5</td>
</tr>
<tr>
<td>Shallice &amp; Evans (1978)*</td>
<td>25</td>
<td></td>
<td></td>
<td>mixed source (age 40-78 yrs., mean 62 yrs.)(mean FSIQ 112)</td>
<td></td>
</tr>
<tr>
<td>Shallice &amp; Burgess (1991)</td>
<td>3.0</td>
<td>2.6</td>
<td>24</td>
<td>no details</td>
<td>8.2</td>
</tr>
<tr>
<td>Shoqirat et al. (1990)**</td>
<td>3.7</td>
<td>3.0</td>
<td>31</td>
<td>matched to patients for IQ (mean FSIQ 103)&amp; age (48 yrs. ± 12) and socio-economic status</td>
<td>9.7</td>
</tr>
<tr>
<td>O'Carroll et al. (1994)**</td>
<td>5.3</td>
<td>3.6</td>
<td>150</td>
<td>general population; age 52 yrs. ± 26 yrs.; mean FSIQ 109</td>
<td>12.5</td>
</tr>
<tr>
<td>Diamond et al. (1996)*</td>
<td>3.7</td>
<td>3.0</td>
<td>6</td>
<td>matched to patients for IQ (FSIQ 103)&amp; age (35 yrs. ± 14)</td>
<td>9.7</td>
</tr>
</tbody>
</table>

* Used Shallice & Evans (1978) version of the CET.
** Used Shoqirat et al. (1990) modified version of the CET.

³ Applying same criteria as Axelrod & Millis (1994) i.e. > 2 sd.
O'Carroll et al. (1994) a similar mean score and standard deviation (4.8; sd. 3.0) were found (Alcott, Stott, Norstrand, Graham, et al; in preparation).

The link between CET performance and the frontal lobes, which has been studied, has theoretical importance. Although Shallice & Evans (1978) suggested that the CET was sensitive to frontal lobe damage or dysfunction, they included patients with lesions involving other areas in addition to the frontal lobes. Therefore, the results may reflect more than frontal dysfunction. In opposition to this argument, they analysed results for patients with damage localised to one lobe, and still found that the degree of bizarreness in responses depended upon the lesion location (i.e. by implication, the frontal lobe). However, Taylor & O'Carroll (1995) did not find the CET to be sensitive to lesion location (anterior vs. posterior) in a study comparing ten neurological conditions (n=257). Because the frontal lobes are the putative site of executive functions this is relevant to the theoretical account.

One possible account for the lack of a consistent correlation between frontal damage and impaired CET performance, is that the CET is sensitive to damage in a specific frontal area. There is some support for this possibility in a study by Leng & Parkin (1988) who compared patients with orbito-frontal damage (i.e. encephalitis and ruptured ACoAA) and those with dorso-lateral frontal damage (i.e. Korsakoff's syndrome). The orbito-frontal patients were impaired on the CET while the dorso-lateral patients were not. However, in "revisiting" Leng & Parkin's study, Shoqirat et al. (1990) found that patients with dorso-lateral damage were as impaired as those with orbito-frontal damage. Resolving this difference is difficult due to differences between the two studies. Leng & Parkin used the original, 15 item CET, while Shoqirat et al. used their modified 10 item version. The sample of Shoqirat et al. is also much larger (i.e. two to five times more patients). As discussed earlier, it is possible that large groups may be needed to obtain a representative sample.
The CET’s specificity for frontal dysfunction can also be questioned. "Abnormal" responses and/or scores are found in “normal” subjects (Axelrod & Millis 1994, Shoqeirat et al. 1990). Using the cut-off score calculated to provide the most accurate differentiation of brain injured patients and non-patients, Axelrod & Millis found that 13% of their “normal” sample would be mis-classified as “impaired”. This suggests that factors other than the integrity of frontal functions may determine CET performance. However, it is possible that such errors occur even in the absence of frontal damage particularly as there is evidence of executive ‘dysfunction’ occurring in people without brain injury (e.g. Reason 1979).

### TABLE 2
Correlation Between the CET and ‘Frontal’ Tests

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sorting task</td>
<td>.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAS</td>
<td>.05</td>
<td>-.40</td>
<td>-.34</td>
<td>-.40</td>
</tr>
<tr>
<td>M-WCST</td>
<td>-.10</td>
<td>-.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>categories</td>
<td>-.10</td>
<td>-.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% perseverative errors</td>
<td>.04</td>
<td>.30</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>WCST</td>
<td>-.25</td>
<td>.24</td>
<td></td>
<td>.25</td>
</tr>
<tr>
<td>categories</td>
<td>-.25</td>
<td>.24</td>
<td></td>
<td>.25</td>
</tr>
<tr>
<td>% perseverative errors</td>
<td>.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>-.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category naming</td>
<td>-.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trails Test (form B)</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Weigl Colour-Form Sort, 2 verbal fluency test, 3 Modified Wisconsin Card Sorting Test, 4 Wisconsin Card Sorting Test, 5 alternating categories

Research comparing the CET with performance on other ‘frontal’ tests is also relevant to the link between frontal lobe function and the CET. Studies by Kopelman (1991), Mayes & Gooding (1989), Shoqeirat et al. (1990) and Wilson, Alderman, Burgess, Emslie & Evans (1996) have failed to find significant correlations between the CET and other standard tests sensitive to frontal dysfunction (see Table 2).

---

4 Correlations are for the patient sample of Korsakoff’s and Alzheimer’s patients.
One more recently developed test of frontal functions has been found to correlate significantly with the CET. The Six Elements Test (SET; Shallice & Burgess 1991b) has been used to assess impaired executive function in patients with frontal lobe damage. The original SET has been modified (M-SET) and incorporated as part of the Behavioural Assessment of Dysexecutive Syndrome (BADS; Wilson, Alderman, Burgess, Emslie & Evans 1996). M-SET performance was compared to CET performance in a study by Burgess & Taylor (Burgess & Shallice 1994; op. cit.). Using a patient sample of mixed aetiology (primarily TBI), they explored performance on various “supervisory variables” contributing to M-SET performance. CET performance was significantly correlated to their measure of “planning” (i.e. the complexity and detail of the intended solution plan). Unfortunately there was no report of the degree of association between the CET and the Temporal Estimation subtest of the BADS.

The association between intelligence and CET performance is important in view of the suggestion that frontal lobe functions underlie performance on tests of fluid intelligence (Duncan, Burgess & Emslie 1995). Although Shallice & Evans (1978) did not find an association between intelligence and CET performance, this has been reported by others. CET performance was found to correlate with current intelligence (Shoqeirat et al. 1990, Kopelman 1991) and premorbid intelligence (Kopelman 1991, O’Carroll et al. 1994). There is also an association between educational achievement and CET performance (Axelrod & Millis 1994, O’Carroll et al. 1994). Lower intelligence and less educational achievement are associated with greater bizarreness on the CET.

**The Frontal Lobes and Estimation: Other Evidence**

An association between impaired estimation and frontal lobe damage was also found by Smith & Milner (1984) who asked patients to estimate the prices of 16 common household items represented by toys. Those with right frontal lobe damage made significantly more errors than patients with left frontal damage,
and patients with temporal lobe damage (regardless of laterality of temporal lesion). They suggest that the absence of a significant impairment in the left frontal patients might be because their lesions tended to be smaller.

The method of scoring errors used by Smith & Milner may have affected the results. Although defining an error rationally (i.e. greater than two standard deviations from the controls’ mean estimate) is more objective than Shallice & Evans’ (1978) method it excludes less deviant responses (i.e. greater than +1 s.d. but less than +2 s.d.). Inclusion of these, as done by Shallice & Evans, might have affected the apparent hemispheric difference.

In another study using a different type of estimation task, Smith & Milner (1988) found an impairment in patients with a frontal lobe lesion, regardless of hemisphere. They presented a series of abstract designs in which each design appeared a specific number of times (with a frequency of 1, 3, 5, 7 or 9) and afterwards asked the subjects how often each design had been presented. The frontal lobe group showed a significantly greater underestimation of the frequency of occurrence than the normal control group. However, this effect depended on the actual frequency with which the designs had appeared in the presentation list. Both right and left frontal lobe groups differed significantly from the control group only at the maximum frequency level (i.e. occurring nine times). At the next lower level of frequency (i.e. seven) only the right frontal lobe group was significantly different from controls and at frequencies below this there was no significant difference between controls and patients.

An important issue with regard to this discussion, is whether Smith & Milner’s task is truly an estimation task. On the basis of the reported mean scores it would appear that at the lower levels of occurrence (e.g. <3) it is not necessarily an estimation task as both normal and brain injured groups’ mean ‘estimates’ appear to be perfectly accurate. These responses may reflect recall of experience rather than an estimate based on recalled experience (for support

see Brown 1995). Without details of the range and distribution of the responses at these frequency levels this is difficult to determine. As they point out in their discussion, the results may indicate a dysfunction of memory (e.g. deficient formation of multiple traces in memory) or of executive function (e.g. strategy of memory search or formulation of estimate response).

In another study of event frequency estimation reported by Brown (1995), university students were asked to estimate how often particular words had occurred throughout the presentation of a lengthy list of word-pairs. They were instructed to try and remember the target word of each pair (e.g. CITY - London) but were subsequently asked to estimate how often specific target words had occurred (i.e. actually 2, 4, 6, 8, 12 or 16 times). In one condition the target words always appeared with the same associated word (e.g. always CITY - London; ‘same-pair’). In another condition the target appeared with various associated words (e.g. CITY - Bath, CITY - Reading, CITY - Bristol, etc.; ‘different pair’). The results support the suggestion that different strategies were used to suit different conditions. In the ‘different-pairs’ condition they tended to use an ‘enumeration-based’ strategy whereas in the ‘same-pairs’ condition the strategy tended to be ‘non-enumeration-based’ (e.g. ‘context-based retrieval’). Enumeration strategies “produce numerical information in the form of counts that can be used to determine an event’s exact or approximate frequency . . . to draw inferences about the response range and to anchor subsequent estimates”. Non-enumerative strategies “provide a qualitative or relative evaluation of event frequency that must be converted to an appropriate numerical value”. The significance of this study to the present review is that it supports Shallice & Evan’s suggestion that estimation is a problem solving process involving a solution plan.

Vilkki & Holst (1991) reported a task that involved people predicting (i.e. estimating) their performance on the Digit Symbol sub-test of the WAIS (i.e. estimating how many items they would complete in one minute). They used
‘adequacy of estimation of sub-goals’ to investigate the effect of frontal damage on ‘mental programming’. Participants were told how many items they had completed in one minute during the initial trial. On several subsequent trials they were asked to estimate how many they would complete. The accuracy of their estimations were measured in terms of the number of items (symbols) completed and the time taken to complete them (if less than the minute allowed). The degree of over-estimation was measured as the average time on trials when the goal was achieved before the minute elapsed (shorter average times indicate greater over-estimation). The degree of under-estimation was measured by the average percentage of symbols that had not been completed when the minute had elapsed before their goal was achieved (higher percentage of uncompleted items indicates greater under-estimation). They found that the patients with frontal damage were less accurate in their estimations and concluded that this indicated a mental programming deficit.

Vilkki & Holst’s measures of ‘adequacy’ of estimation are somewhat unusual. By allowing patients to make their estimates after they have observed their performance on the first trial they differ from most other estimation tasks which do not provide feedback on performance during or before making an estimation. It would have been interesting to obtain a ‘blind’ estimate before the first trial as this might have elicited the type of ‘bizarre’ responses found by Shallice & Evans. A between group comparison of the degree of improvement in accuracy of estimation over subsequent trials would also have been interesting. In particular it would have been interesting to know whether the frontal patients’ estimations improved to the same degree as the controls.

Vilkki & Holst used a ‘risk-taking’ format which may have influenced estimation behaviour and therefore raises questions about their conclusions. Subjects were penalised if their performance fell short of their estimated goal. This appears to have produced a tendency towards conservative estimation. The average time used to reach estimated goals before the minute expired was

significantly shorter in the frontal group compared to the controls. Hence their results may indicate more about the effect of frontal damage on willingness to take risk. Their inaccuracy of estimation may have been due to being less willing to take risk and consequently tending to underestimate.

Taylor (1993) studied the relationship in patients with MS between their performance on cognitive tests, including the CET, and their cognitive difficulties in everyday life. He found that patients with impaired estimation on the CET tended also to be poor at estimating the frequency with which their symptoms and disabilities affected them in everyday life (i.e. underestimation). Their CET score correlated significantly \( r = -0.44 \) with the degree to which they underestimated the frequency of their problems. The weakness with the method is the use of relative’s rating as the criteria for accuracy of estimation which is of questionable reliability.

The ability to estimate time has been studied in various forms in both brain damaged (BD) and non-brain damaged individuals. Non-BD individuals show a characteristic pattern of inaccuracy in estimating how long ago an event occurred (e.g. a news item). They tend to over-estimate shorter intervals (i.e. more recent events) and under-estimate longer intervals (Ferguson & Martin 1983). This tendency results in an overall level of accuracy which is relatively poor. For example, when estimated time was regressed against actual time (using log-linear conversion) the co-efficient of regression was 0.66. Despite such inaccuracy in this study, the high correlation between actual and estimated time \( r = 0.97 \) suggests they did not produce many ‘bizarre’ estimates.

Different forms of time estimation task were used by Grafman, Litvan, Gomez, & Chase (1990) in a study of executive function and attention with PSP\(^5\) patients. Participants were asked to estimate the correct time, elapsed session duration and duration of a short period (60 seconds) during which they counted

\(^5\) Progressive Supranuclear Palsy.

the number of taps presented (6, 15, 30, 60 taps). Patients performed significantly worse than the controls on the tests of executive function (e.g. WCST, FAS, Trails B and Similarities sub-test of the WAIS-R). However, they were not significantly different from the control group on the time estimation measures. Unfortunately there was no report of any analysis of the relationship between executive function tests and time estimation measures. As PSP is associated with psycho-motor slowing (Lees 1990) which could affect the perception of the passage of time it suggests that this is not influencing estimation or they are compensating for it.

Estimation of time has also been used to measure and study the ‘dysexecutive syndrome’ (DES) associated with frontal damage. In an unpublished study, Alderman (1994) reported the use of two time estimation tasks. One involved estimating the passage of real-time (i.e. periods of 10, 30 and 60 second duration) and the other required estimating the duration of hypothetical activities or events (e.g. “How long does it take to boil an egg?”; “How long do most dogs live for?”). He found that ability to estimate real-time did not distinguish the frontal group from the control group. However, performance on the time estimation questions did distinguish the groups. This task also correlated significantly with a questionnaire measure of the severity of DES (i.e. an informant’s rating of severity of various DES symptoms). It did not however correlate significantly (r= 0.14) with the measure of insight impairment (i.e. the discrepancy between the patient’s and informant’s ratings).

The ability to estimate a country’s land area or population size was studied Brown & Seigler (1993). In a series of experiments they focused on the effect on estimation accuracy of providing useful supplementary information. This was either information about the relative, ordinal population size of several similar countries (i.e. which were more or less populous; called ‘mapping’ information) or information about the actual population of related countries (called ‘metric’ information). They found that the psychology students in their

American university were quite inaccurate in their estimations, which they suggest reflects in part their lack of knowledge about the rest of the world. They were however, able to use supplementary information to improve their accuracy according to the type of information provided. Metric accuracy improved with metric, not mapping information and vice versa. The greatest improvement resulted when both metric and mapping information was available. This study is of particular interest for several reasons. Firstly, the degree of inaccuracy in estimating raises the question of how valid or useful estimation tasks can be for distinguishing people with brain injury. Secondly, it supports the importance of semantic memory in estimation that was suggested by Taylor & O'Carroll (1995). Thirdly, and most importantly, it supports the proposed role of executive functions in estimation. The students are clearly able to monitor and regulate their estimation behaviour in order to make use of the supplementary information.

Summary
Since it was devised in 1978 by Shallice & Evans as a measure to explore the effects of frontal lobe damage, the CET has been modified and an Americanised version has been produced with more rationally based scoring criteria. Although the CET is commonly used in neuropsychological assessment in the UK and has been used in many clinical studies as a measure of frontal lobe dysfunction, there has not been any published research exploring the theoretical account proposed by Shallice and colleagues. These accounts suggest that CET performance reflects executive functions of planning, selection and regulation as well as semantic memory access. Although CET performance is not associated with performance on memory tests (i.e. new learning) the possibility of an association with impaired semantic memory has not yet been investigated.

Although the CET is purported to be sensitive to the integrity of the frontal, executive functions it does not correlate significantly with other ‘frontal’ tests. This does not necessarily mean that it does not reflect frontal functioning as
there appear to be a variety of executive functions producing dissociations between various ‘frontal’ tests (Burgess & Shallice 1994). Although abnormal performance is also not restricted to frontal damage this may be due to the fact that dysfunction can occur with lesions in other areas (e.g. subcortical white matter). Whether performance depends upon dysfunction in a specific area of the frontal lobes is not yet clear.

The existence of ‘abnormal’ scores in people without a brain injury and the variation in normative data raises a question about the specificity of the CET to frontal damage and dysfunction. However, dysexecutive-type behaviour has been found in ‘normal’ people and normative variation may reflect insufficient sample sizes or item heterogeneity. Although the latter could weaken the CET’s psychometric characteristics, it might also be essential to require repeated planning during the task. This would be important, if as suggested by Shallice & Evans, the CET reflects executive functions such as strategy planning.

Although there are a variety of other estimation studies suggesting that frontal functions are involved, most have some limitation that raises the question of their applicability to the issue. The estimation studies with non-brain injured people provides support for the argument that estimation involves strategy planning and the capacity to monitor and regulate behaviour.

**Testing the Explanations of CET performance**

A test of the dysexecutive explanations of CET performance would need to involve controls for planning, strategy selection and regulation. Testing the memory account would need to involve varying access to knowledge (i.e. semantic memory). One complication that would affect a test of both accounts is the issue of differences between people in their level of relevant knowledge.
This could be circumvented by providing relevant information to equalise peoples' knowledge base. This is essentially the effect of the method used by Brown & Seigler (1993). However, the CET differs from the tasks they used which focused on specific topics (e.g. country's populations) in that the CET involves a variety of different knowledge domains (e.g. height of buildings, human anatomy, weight of objects, etc.). Because of this it would be necessary to provide a different set of information for each CET item. Providing information in this way is likely to be more salient than providing a single set of information, therefore it could act as a prompt to use the information. This could activate the Supervisory System in the problem solving process with the result that an executive deficit would be masked (see Shallice & Burgess 1991b). As a result, directly applying Brown & Seigler's method to test accounts of the CET presents a potential complication for interpreting the results should there be no difference between those with frontal damage and those without (i.e. risk of a possible Type II Error).

In view of this another estimation task was designed which fitted the following specifications: (a) the true value of each item would be discrete and 'knowable' (to allow accuracy to be quantified), (b) the exact value was unlikely to be known by the subject (thus necessitating an estimation), (c) all items would relate to a single general domain of knowledge (to allow a single set of supplementary information to apply to all items), (d) the items would all fall along a single, continuous dimension (so that supplementary information could provide some clue about the actual value of each item as well as information about its relative position on the dimension), (e) the range of this dimension should be broad (to allow selection of a sufficient number of items and such that there was potentially a wide range of error), (f) the domain of knowledge is one which is likely to be familiar to most adults (to aid potential for possible future clinical use as a measure applicable to most people).

The task selected on the basis of these criteria was that involving estimation of distance between pairs of well known towns and cities in mainland Britain. As it requires the estimation of mileages it was named the “Milest Test”. It consists of 40 pairs of towns/cities (e.g. London - Brighton) whose inter-city distance is to be estimated.

The Milest Test was designed in two forms, one with and one without additional information. The additional, or ‘supplementary’ information combines both ‘metric’ and ‘mapping’ information. This combination has the greatest potential for benefiting accuracy of estimation (Brown & Seigler 1993). The mapping information (i.e. information about ordinal relationships between the various inter-city distance) was a small scale map of mainland Britain showing the location of all the relevant towns and cities. The metric information was the mileage between three pairs of towns which they were not required to estimate.

**Aims**

The present study was designed to examine the theoretical account of the CET. The main aim is to address the hypotheses that CET performance reflects the integrity of executive, supervisory functions, particularly monitoring and regulation of the outcome or possibly arises from deficient access to or retrieval from semantic memory. A secondary aim of the study is to increase our understanding of the psychological meaning of performance on estimation tests. At present, clinical interpretation of CET performance is tenuous as it rests on an assumption that the proposed theoretical account is valid. With a clearer understanding of the psychological processes involved in CET performance it would not be necessary to rely on such an assumption. Also it would hopefully be easier to extrapolate to real-life situations based on knowledge of the psychological processes which are impaired.

The Milest test, described above, was chosen for these purposes. The executive dysfunction hypothesis would predict that patients with a brain injury affecting executive functions should be less able to use supplementary information when completing the Milest test, hence they would not improve as much as people without a brain injury.

The importance of semantic memory access is also addressed using this method. Providing supplementary information would negate the need to access semantic memory. Therefore this hypothesis would predict that when such information is available there would be less difference between those with and those without frontal damage.

METHOD
PARTICIPANTS
There were four groups of participants; Experimental BI group (46 men, 12 women; mean age = 45.43 yr., sd = 13.94), Experimental Non-BI group (28 men, 18 women; mean age = 42.17 yr., sd = 13.97), Control BI group (18 men, 10 women; mean age = 36.43 yr., sd = 13.94) and Control Non-BI group (17 men, 26 women, 3 unknown; mean age = 34.78 yr., sd = 11.50). Participants’ level of educational achievement is shown in table 3.

<table>
<thead>
<tr>
<th></th>
<th>BI Experimental</th>
<th>BI Control</th>
<th>Non-BI Experimental</th>
<th>Non-BI Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>No exams</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>GCSE/O'levels</td>
<td>18</td>
<td>14</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>A'levels</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>University</td>
<td>14</td>
<td>3</td>
<td>15</td>
<td>25</td>
</tr>
</tbody>
</table>

6 Details of approximate age were collected rather than exact age (i.e. aged 16-25 yr., 24-35 yr., etc.) therefore mean age was calculated using the number of people in each category multiplied by the category median.

Participants in the BI groups had an acquired brain injury. In the Experimental and Control groups the diagnoses were respectively, traumatic brain injury (18 and 18), cerebral haemorrhage (17 and 3), cerebral tumour (1 and 1), encephalitis (2 and 1), anoxia (2 and 1) and other (3 and 2). Patients were included in the study where there was likely frontal lobe damage or dysfunction on the basis of the nature of injury (e.g. rapid deceleration head injury), the results of brain imaging scans, or symptoms (e.g. behavioural disinhibition). Patients were excluded if other clinical factors were likely to influence their performance (e.g. dyslexia, dysgraphia, severely impaired comprehension, confusion or impaired visuo-spatial perception).

The participants in the Non-BI group were all adults without acquired brain injury. They were from various sources (hospital staff, orthopaedic patients, students, neighbours, etc.). The Non-BI participants were selected from a larger Non-BI sample (n=106) in order to achieve the best possible match with the BI participants in terms of age and level of education. Selection aimed to achieve the best possible balance according to various independent variables. Therefore selection was made with regard to sex, age and education but without reference to participants’ Milest performance. Some Non-BI participants were literally matched to the BI participants (i.e. same sex, age and education) but as this was not possible in all cases the over riding aim was to achieve a similarity between the groups.

The groups were compared using a non-parametric test for interval data (i.e. Kruskal-Wallis one-way ANOVA) for the variables of age and educational achievement which were converted from categorical to ordinal data. A chi-square test was used to compare the dichotomous data (i.e. sex). All four groups were included in the comparisons. There were significant differences between the groups concerning age (chi-square = 17.13, d.f.= 3, p<.001), education (chi-square = 14.27, d.f.= 3, p<.01) and sex (Pearson chi-square =
11.38, d.f.=1,3, p<.01). In view of this subsequent analyses were recalculated, where appropriate, using these variables as co-variates.

MEASURES

The Milest test described above was the experimental measure. The CET used was the modified form of Shoqeirat et al (1990).

PROCEDURE

All participants were given the Milest Test form and asked to write what they thought to be the distance between each of the 40 pairs of towns. They were told that they were likely to need to estimate the distances, as it was unlikely that they would know any or many of the actual distances. On the first occasion all the participants completed the Milest without supplementary information (i.e. map with the distance information). On this occasion they were instructed not to refer to any other information (particularly the latter part of the form).

Immediately thereafter they completed the Milest a second time. The version used (i.e. with or without supplementary information) on the second occasion depended upon whether they were in the Experimental or Control group. Participants were not allowed to refer back to their initial estimations on the second occasion. The CET was also completed by the participants.

Several measures were drawn from the Milest including error/accuracy scores and indicators of dysexecutive behaviour. The primary measure of accuracy was the proportional error score. This was calculated for each item by finding the absolute error and dividing it by the actual distance for each item. For example if the estimated mileage was 60 miles when the actual distance was 80 miles then the proportional error was 0.25 (i.e. \(|80-60| ÷ 80 = 0.25\)). A total
proportional error score was calculated for each participant on both the first and second administration of the Milest Test (i.e. Estimate 1 and Estimate 2).

Accuracy was also measured using the results of the regression of actual distances against the estimated distances. This produces three inter-related measures; the squared correlation between actual and estimated distance ($r^2$), the linear slope ($b$) and the intercept (‘constant’). Each reflects different aspects of estimation accuracy which may help to elucidate any inter-group differences.

If there was a perfect relationship between estimated and actual distance then $r^2$ and $b$ would equal one and the constant would be zero (see example B in figure 1). The slope ($b$) is analogous to ‘gain’ in that it reflects whether a small change in the independent variable (x axis) results in a large or small change in the dependent variable (y axis). In figure 1 two types of slope are shown, with $b$ greater than one (example A) or less than one (example C). If actual distance were represented by the x axis then example A indicates that any discrepancy between actual and estimated distance is greater for longer distances than it is for shorter distances. Example C would indicate that any discrepancy between actual and estimated distance is less for longer distances than it is for shorter distances.

---

7 In the linear regression the dependent variable was the actual mileage.

The intercept is analogous to a constant error represents a constant correction value that needs to added or subtracted when predicting the response from the actual value. In this case it reflects whether there is a tendency to over or under estimate. It also indicates the degree of this tendency. Two examples of constants are shown in figure 2, one greater than one (example A) and one less than one (example C). A tendency to over estimate is reflected in example A while example C reflects a tendency to under estimation.

The constant is important to the interpretation of the slope. For example, if the constant were greater than one then a slope greater than one would indicate that the degree of error increases as the distance increases. If the constant were less than one the same slope would indicate that the degree of error decreases with increasing distance. Consequently if the constant is greater than one then a value of $b$ greater than one would it indicate that there is a tendency to over-estimate which increases for longer distances than for shorter distances. For the same constant a $b$ less than one would indicate that the degree of error tends to be less for longer distances.

The correlation squared ($r^2$) is analogous to the extent of ‘coupling’ between actual and estimated distances. If reflects the degree of consistency in the relationship between the estimated and actual distance. The higher the value of

---

the more consistent, and therefore the more predictable the relationship between the estimates and the actual distances.

The dysexecutive measures extracted from the Milest were perseverative and bizarre responses. Perseveration is a measure that has been used previously, particularly to explore cognitive dysfunction associated with frontal dysfunction (e.g. the WCST; Grant & Berg 1948). Two types of response pattern were used as measures of perseveration; repetitive and modal responses. Repetitive perseveration was defined as repetition of the preceding response. When the same estimation response was repeated more than twice in succession each response after the first one in the series was counted as an instance of perseveration. Modal perseveration was defined as the tendency to use a particular response repeatedly throughout the task (i.e. not necessarily successive responses). The modal response was determined for each participant for each administration of the Milest Test (i.e. a separate modal response was found for the first and second administration occasions). The total number of modal responses was calculated for each occasion.

Determination of the criteria for bizarre responses was done on the same basis as that used by Axelrod & Millis (1994). Following this reasoning, bizarre responses were defined as those that were at least two standard deviations from the mean response of the Experimental Non-BI group. Separate criteria were calculated for each of the 40 Milest items for the first and second administration occasions. The total number of bizarre responses was then calculated for the first and second estimate separately.
DESIGN

A mixed design with repeated measures between groups was used. There were two between-subject factors, Brain Status (with or without a brain injury) and Group (Experimental or Control). The repeated measure was Occasion (first and second administration of the Milest test). This is summarised in table 4.

<table>
<thead>
<tr>
<th>Group</th>
<th>First Occasion</th>
<th>Second Occasion</th>
<th>(CET)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI (E) Without Information</td>
<td>With Information</td>
<td>(CET)</td>
<td></td>
</tr>
<tr>
<td>Non-BI (E) Without Information</td>
<td>With Information</td>
<td>(CET)</td>
<td></td>
</tr>
<tr>
<td>BI (C) Without Information</td>
<td>Without Information</td>
<td>(CET)</td>
<td></td>
</tr>
<tr>
<td>Non-BI (C) Without Information</td>
<td>Without Information</td>
<td>(CET)</td>
<td></td>
</tr>
</tbody>
</table>

BI = Brain Injured  Non-BI = Non-Brain Injured  E = Experimental  C = Control

Participants completed the Milest twice (First and Second Occasion). On the first occasion all the participants made their estimates without supplementary information. On the second occasion those participants in the experimental condition completed the Milest Test with supplementary information. Those in the control condition completed the Milest on the second occasion without supplementary information.

SPECIFIC PREDICTIONS AND STATISTICAL ANALYSIS

Based on the theoretical proposals, various predictions can be made about performance on the Milest Test by people with a brain injury (BI group) compared to those without a brain injury (non-BI group):

1. The brain injured participants will be worse at estimating. Therefore the BI groups will have a significantly higher total proportional error score than the Non-BI groups on the first occasion. This will be tested using a t-test to compare the total proportional error scores on the first occasion of the individuals in the combined BI and Non-BI groups.

2. The brain injured participants' poorer accuracy will also be reflected in the value of the linear regression constant. Therefore, the BI group's mean constant value will be a significantly larger than the Non-BI group's on both occasions. This will also be tested as in the first prediction using a t-test to compare the combined BI and Non-BI groups.

3. The BI group will also be less consistent in their estimation performance. Therefore the BI group's values of $r^2$ will be significantly lower than the Non-BI group's. The BI group's $b$ value will also be significantly lower than the Non-BI group's. These comparisons will also be made using a t-test to compare the combined BI and Non-BI groups on the first occasion for the values of $r^2$ and $b$.

4. The brain injured participants will benefit less from the supplementary information hence they will improve their accuracy of estimation less than the non-brain injured participants on the second occasion. Therefore there will be a significant three-way interaction between Brain Injury Status, Occasion and Group. The Experimental BI group will improve less than the Experimental Non-BI group on the second occasion. If this is due to the use of the supplementary information rather than practice then the BI and Non-BI Control groups will not significantly improve on the second occasion. This will be tested using a mixed analysis of variance (ANOVA) for repeated measures between groups to compare their total proportional error scores and their linear regression measures.

A significant three-way interaction between Brain Injury Status, Occasion and Group could also support the semantic memory deficit hypothesis. However, there would be a different reason for the interaction. Instead of improving less than the Non-BI group on the second occasion, the Experimental BI group would improve to a greater extent. This too will be tested using a mixed ANOVA for repeated measures followed by post hoc multiple comparisons test to examine whether the BI group improved to a greater or lessor extent that the Non-BI group on the second occasion. Tukey’s HSD will be used for post-hoc comparisons as recommended by Kinnear & Gray (1995).

5. There will be evidence of dysexecutive behaviour in the brain injured participants. Therefore the BI group will have significantly higher CET scores, perseveration scores and make significantly more bizarre responses than the Non-BI group. This will be tested using t-tests to compare the two groups using these measures from the first occasion.

6. Although the supplementary information should reduce the degree of any dysexecutive behaviour in the Experimental Non-BI group, it should have less impact on the Experimental BI group further to the fourth prediction. Therefore there should be a significant three-way interaction between Brain Status, Occasion and Group with the Experimental BI group improving less than the Experimental Non-BI group on the second occasion in the analyses of perseverative and bizarre responses. Mixed ANOVA with repeated measures between groups will used to test this with the measures of perseveration and bizarre responses.

7. Because the CET measures bizarre (i.e. “extreme”) responses there should be a significant positive correlation between it and the measure of bizarre responses on the Milest Test on the first occasion without supplementary information. Spearman’s non-parametric correlation will used to measure
the degree of correlation between total CET and total bizarre responses on the first occasion in the combined groups (Experimental and Control BI and Non-BI groups).

The relationship between bizarre responding and other measures of estimation behaviour will also be explored. This will done by analysing the correlations in both the BI and Non-BI groups between the other Milest measures on the first occasion as well as the CET.

RESULTS

Effect of Brain Injury on Accuracy of Estimation

To compare the accuracy of estimation between the BI and Non-BI participants the Experimental and Control groups were combined to form BI and Non-BI groups. Several measures of estimation accuracy were extracted from the First Occasion (i.e. Time 1) and then used to compare the two groups. These included total proportional error score (PE), the linear regression intercept value (constant), the linear regression correlation squared ($r^2$) and the slope of the linear regression line ($b$). Levene’s Test for equality of variance was used. Because it was significant in all analyses the two-tailed significant level for unequal variances is reported.

Total Proportional Error

There was a significant difference ($t=6.99; p<.009$) in the mean total PE of the two groups. The BI group’s mean total PE score (23.0) was significantly higher than the Non-BI group’s (14.9).

Linear Regression Intercept

The mean value of the linear regression intercept was not significantly different between the BI and Non-BI groups ($t=0.03; p=.98$). The BI group’s mean was 21.2 compared to the Non-BI group’s mean of 21.1.
**Linear Regression Correlation Squared ($r^2$)**

There was a significant difference ($t = -9.08; p = .03$) in the mean $r^2$ value of the two groups. The BI group’s mean $r^2$ value (.49) was significantly lower than the Non-BI group’s (.76).

**Linear Regression Slope ($b$)**

There was a significant difference ($t = 2.75; p = .007$) in the mean $b$ value of the two groups. The BI group’s mean $b$ value (1.22) was significantly higher than the Non-BI group’s (0.98).

**Summary:** The BI participants were significantly less accurate at estimating inter-city distance on the measure of total proportional error. They were also less consistent and indicated by the analyses of $r^2$ and $b$. However, there was no difference on the linear regression intercept, which suggests that the BI participants do not show an atypical tendency in over-estimation.

**Ability to Benefit from Supplementary Information**

The Milest measures (total PE and $r^2$, $b$ and intercept from the linear regression correlation) were analysed using a mixed ANOVA for repeated measures. Separate ANOVA’s were done with the repeated measure being Occasion (i.e. First and Second administration of the Milest). The two between-subjects factors were Group (i.e. Experimental and Control) and Brain Status (i.e. With Brain Injury and Without Brain Injury). All analyses were repeated, entering Education, Age and Sex simultaneously as covariates. The primary result of relevance to the predictions is the three-way interaction as this would indicate whether there is a difference between the BI and Non-BI groups and whether this is due to information, rather than practice. Post-hoc comparison of mean scores from the ANOVA’s was done using Tukey’s HSD. This test

---

8 Categorical data was converted to ordinal data.
9 In calculating the Critical Difference threshold the level of probability was $p < .01$.

compensates for increases in the number of comparisons by adjusting the level of ‘Critical Difference’ between means required to achieve significance.

Total Proportional Error

There was not a significant interaction between Brain Status, Occasion and Group (F= 0.09, d.f.= 1,162, p=.76). The means for the groups are shown in table 5 and these are plotted in figure 3.

Table 5
Mean scores relating to the ANOVA of Proportional Error scores

<table>
<thead>
<tr>
<th>Three-way Interaction</th>
<th>Brain x Group x Occasion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental BI Group</td>
<td>Control</td>
</tr>
<tr>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td>Experimental</td>
<td>22.62</td>
</tr>
<tr>
<td>Non-BI Group</td>
<td>13.30</td>
</tr>
<tr>
<td>Control</td>
<td>(8.77)</td>
</tr>
<tr>
<td>Control</td>
<td>(6.33)</td>
</tr>
</tbody>
</table>

The two BI groups had a higher mean total PE score on both occasions compared to the Non-BI groups. Only the Experimental groups improved on the second occasion. The Experimental BI group improved to the same degree as the Experimental Non-BI group.

![Figure 3](image)

Figure 3
Mean Total PE: Brain Status by Group by Occasion

The ANCOVA using Education, Age and Sex as covariates did not alter the level of significance for the three-way interaction \( (F= 0.13, \text{ d.f.}= 1,162, p=.72) \).

*Linear Regression Intercept*

The intercept ('constant') from the linear regression was also analysed using a mixed ANOVA for repeated measures. The mean intercept values relating to three-way ANOVA are shown in table 6.

Table 6

Mean scores relating to the three-way ANOVA of linear regression intercept

<table>
<thead>
<tr>
<th>Brain x Group x Occasion</th>
<th>BI Group</th>
<th>Non-BI Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experimental</td>
<td>Control</td>
</tr>
<tr>
<td></td>
<td>1st 2nd</td>
<td>1st 2nd</td>
</tr>
<tr>
<td></td>
<td>23.66 12.51</td>
<td>17.27 9.51</td>
</tr>
<tr>
<td></td>
<td>(34.07) (23.12)</td>
<td>(43.33) (63.62)</td>
</tr>
</tbody>
</table>

There was not a significant interaction between Brain Status, Occasion and Group \( (F= 0.4, \text{ d.f.}= 1,162, p=.53) \). The ANCOVA using Education, Age and Sex as covariates did not alter the level of significance for the three-way interaction \( (F= 0.13, \text{ d.f.}= 1,162, p=.72) \). The results are shown in figure 4.

![Figure 4](image)

**Figure 4**

Mean intercept value: Brain Status by Group by Occasion

On the first occasion there was no significant difference between any of the four groups. Both BI and Non-BI Experimental groups had a lower mean...
intercept value on the second occasion, compared to the Non-BI Control group. The BI Control group also significantly reduce their mean intercept value on the second occasion.

*Linear Regression Correlation Squared*

The $r^2$ value from the linear regression was also analysed using a mixed ANOVA for repeated measures. The mean $r^2$ values relating to three-way ANOVA are shown in table 7.

<table>
<thead>
<tr>
<th>Three-way Interaction</th>
<th>Brain x Group x Occasion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experimental</td>
</tr>
<tr>
<td></td>
<td>1st 2nd</td>
</tr>
<tr>
<td>BI Group</td>
<td>0.54 0.81</td>
</tr>
<tr>
<td></td>
<td>(0.22) (0.14)</td>
</tr>
</tbody>
</table>

There was not a significant interaction between Brain Status, Occasion and Group ($F=2.67$, d.f.= 1,162, p=.10). The ANCOVA using Education, Age and Sex as covariates did not alter the level of significance for the three-way interaction ($F=2.53$, d.f.= 1,162, p=.11). The results are shown in figure 5.

Figure 5

Mean value of $r^2$: Brain Status by Group by Occasion

On the first occasion the Non-BI groups’ mean $r^2$ was nearer to one than the BI groups’. Only the Experimental groups improved on the second occasion, with the Experimental BI group benefiting as much from the supplementary information as the Experimental Non-BI group.

**Linear Regression Slope ‘b’**

The mean $b$ values relating to three-way ANOVA are shown in table 8. The slope of the linear regression ($b$) was also analysed using a mixed ANOVA for repeated measures.

<table>
<thead>
<tr>
<th>Three-way Interaction</th>
<th>Brain x Group x Occasion</th>
<th>BI Group</th>
<th>Control</th>
<th>Non-BI Group</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1st</td>
<td>2nd</td>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.05</td>
<td>1.16</td>
<td>1.49</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.49)</td>
<td>(0.45)</td>
<td>(0.84)</td>
<td>(1.40)</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1st</td>
<td>2nd</td>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.96</td>
<td>0.99</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.25)</td>
<td>(0.10)</td>
<td>(0.35)</td>
<td>(0.40)</td>
</tr>
</tbody>
</table>

There was not a significant interaction between Brain Status, Occasion and Group ($F= 0.16$, d.f.$= 1,162$, $p=.69$). The ANCOVA using Education, Age and Sex as covariates did not alter the level of significance for the three-way interaction ($F= 0.15$, d.f.$= 1,162$, $p=.70$). The results are shown in figure 6.

![Graph](image_url)  
**Figure 6**  
Mean value of $b$: Brain Status by Group by Occasion

On the first occasion only the BI Control is significantly different from the other groups. There is no significant difference between the two occasions in the Non-BI groups. However, there is a significant difference in the two BI groups between the first and second occasion. Overall most of the groups’ $b$ values were very near to a value of one, with the exception of the BI Control group.

Summary: The BI participants who received supplementary information were able to improve the accuracy of estimation to a similar degree as the Non-BI participants. This was the case on all the measures reported. Participants who did not receive supplementary information did not improve their accuracy. The exception was the Control BI group whose mean intercept value did improve.

Effect of Brain Injury on Dysexecutive Measures

To compare the BI and Non-BI participants’ performance on dysexecutive measures the Experimental and Control groups were combined to form BI and Non-BI groups. Several dysexecutive measures were extracted from the First Occasion (i.e. Time 1) and then analysed including total CET score, total repetition perseverations, total modal perseverations and total bizarre responses. Because Levene’s Test for equality of variance was significant for the CET and bizarre responses measures the two-tailed significant level for unequal variances are reported for these.

There was a significant difference in the total CET scores between the BI and Non-BI groups ($t = 5.21; p<.009$). The BI group had a significantly higher mean total CET score than the Non-BI group (means respectively 7.64 and 2.80). There was also a significant difference in the Milest total bizarre response scores between the BI and Non-BI groups ($t = 2.08; p=.04$). The BI group’s mean bizarre response score was significantly greater than the Non-BI
group's (3.2 versus 1.9). The range of bizarre responding the BI group was 0-25 compared to the Non-BI group's range of 0-10.

There were no significant differences between the groups in their levels of perseverative responding. The difference in the mean number of repetition perseverations made by the BI group (2.7) and the Non-BI group was not significant (1.9; \( t = 1.83; p= .07 \)). The mean number of modal perseverations made by the BI and Non-BI groups were virtually the same (6.8 and 6.0; \( t = 1.87; p= .06 \)).

**Summary:** There was a difference between the BI and Non-BI groups in terms of bizarre responding but not on the measures of perseverative responding.

**Effect of Supplementary Information on Dysexecutive Measures**

The effect of supplementary information on estimation for participants with and without a BI was analysed using a mixed ANOVA for repeated measures. Separate ANOVA's were done using the same within and between subjects factors mentioned above. All analyses were repeated, entering Education, Age and Sex simultaneously as covariates. As above, the primary result of relevance to the predictions is the three-way interaction and Tukey's HSD was used to analyse post-hoc comparisons of mean scores and where significant differences where found these were reported.

**Repetition Perseveration**

There was not a significant interaction between Brain Status, Occasion and Group (\( F = 1.96, d.f. = 1,162, p=.16 \)). The ANCOVA using Education, Age and Sex as covariates did not alter the level of significance for the three-way interaction (\( F = 1.78, d.f. = 1,162, p=.18 \)). The means for the groups are shown in table 9 and these are plotted in figure 7.
Table 9
Mean scores relating to the ANOVA of repetition perseverations

<table>
<thead>
<tr>
<th></th>
<th>BI Group</th>
<th>Control</th>
<th>Non-BI Group</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st</td>
<td>2nd</td>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td>Experimental</td>
<td>2.59</td>
<td>3.46</td>
<td>3.00</td>
<td>3.11</td>
</tr>
<tr>
<td>Control</td>
<td>(3.26)</td>
<td>(2.54)</td>
<td>(3.50)</td>
<td>(3.28)</td>
</tr>
</tbody>
</table>

The BI groups had significantly higher mean scores than the Non-BI groups on both occasions. None of the groups made fewer perseverations on the second occasion, although the Experimental BI group's mean score was significantly higher than it was on the first occasion.

![Figure 7](image.png)

Figure 7
Mean Repetition Perseverations: Brain Status by Group by Occasion

**Modal Perseveration**

There was a significant interaction between Brain Status, Occasion and Group (F= 4.26, d.f.= 1,162, p=.04). Similar results were found in the ANCOVA using Education, Age and Sex as covariates (F= 4.92, d.f.= 1,162, p=.03). The means for the groups are shown in table 10 and these are plotted in figure 8.
Table 10
Mean scores relating to the ANOVA of modal perseverations

<table>
<thead>
<tr>
<th>BI Group</th>
<th>Experimental</th>
<th>1st</th>
<th>2nd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.22)</td>
<td>(2.67)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-BI Group</th>
<th>Experimental</th>
<th>1st</th>
<th>2nd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.27)</td>
<td>(1.56)</td>
</tr>
</tbody>
</table>

Figure 8
Mean total modal perseverations: Brain Status by Group by Occasion

On the first occasion the mean scores were similar for all the groups with the exception of the BI Control group. Only the Experimental Non-BI group showed a reduction in their mean score on the second occasion. Both Control groups showed a significant increase in perseveration on the second occasion.

Bizarre Responses
There was no significant interaction between Brain Status, Occasion and Group (F= 0.42, d.f.= 1,162, p=.52). The ANCOVA using Education, Age and Sex as covariates did not alter the level of significance for the three-way interaction (F= 0.5, d.f.= 1,162, p=.48). The means for the groups are shown in table 11 and these are plotted in figure 9.

Table 11
Mean scores relating to the ANOVA of bizarre responses

<table>
<thead>
<tr>
<th>BI Group</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td>3.76</td>
<td>3.39</td>
</tr>
<tr>
<td>(4.81)</td>
<td>(5.27)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-BI Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
</tr>
<tr>
<td>1st</td>
</tr>
<tr>
<td>1.87</td>
</tr>
<tr>
<td>(2.40)</td>
</tr>
</tbody>
</table>

The ANCOVA using Education, Age and Sex as covariates did not alter the level of significance for the three-way interaction ($F= 0.5$, d.f. = 1,162, $p=.48$).

Figure 9
Mean bizarre responses: Brain Status by Group by Occasion

On the first occasion only the Experimental BI group made significantly more bizarre responses than the Non-BI groups. On the second occasion only the Experimental Non-BI group made significantly fewer bizarre responses compared to the first occasion.

Summary: The Non-BI participants who received supplementary information did reduce their level of dysexecutive behaviour on the measures of modal perseveration and bizarre responding. None of the groups improved on the measure of repetitive perseveration. The BI participants who received supplementary information did not show any improvement on the measures of dysexecutive behaviour.

Relationships Between Measures

The relationships between the CET and the Milest measures was analysed using Spearman’s correlation firstly using the performance of all groups (n= 166) on the first occasion. The results are shown in table 12.

Table 12
Correlations between Milest Measures in the four groups combined on the first occasion.

<table>
<thead>
<tr>
<th>PE</th>
<th>.39*</th>
<th>.43*</th>
<th>-.76*</th>
<th>.11</th>
<th>-.03</th>
<th>-.04</th>
</tr>
</thead>
<tbody>
<tr>
<td>r²</td>
<td></td>
<td>-.24*</td>
<td>-.65*</td>
<td>.06</td>
<td>.25*</td>
<td>-.24*</td>
</tr>
<tr>
<td>Constant</td>
<td>.01</td>
<td>.03</td>
<td>-.06</td>
<td>.15</td>
<td>.12</td>
<td>.67*</td>
</tr>
<tr>
<td>Mod. Per.</td>
<td>.07</td>
<td>.28*</td>
<td>-.23*</td>
<td>-.66*</td>
<td>.18*</td>
<td>-.05</td>
</tr>
<tr>
<td>Rpt. Per.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biz. Resp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CET- Cognitive Estimates Test, PE= Total proportional error, r²= linear regression correlation squared, b= linear regression slope, Constant= linear regression intercept, Mod Per= Modal perseverations, Rpt Per= Repetition perseverations

In the four combined groups the CET correlated significantly with the Milest measures of PE and r² but not with bizarre responding or perseveration. The strongest correlations were between PE and r² (r= -.76), linear regression slope (b) and constant (r= -.65), bizarre responding and b (r= -.66) and the two perseveration measures (r= .67). The only measure which correlated significantly with bizarre responding was b.

The same correlations were then computed for the BI and Non-BI groups separately (N.B. In each group the data from the Experimental and Control groups on the first occasion were combined.) The results for the BI group (n=74) are shown in table 13.

Table 13
Correlations between Milest Measures in the combined BI group.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PE</td>
<td>.19</td>
<td>- .39</td>
<td>- .38</td>
<td>- .25 *</td>
<td>- .10</td>
<td>- .14</td>
<td>- .23</td>
</tr>
<tr>
<td>r²</td>
<td></td>
<td></td>
<td></td>
<td>- .03</td>
<td>.30 *</td>
<td>- .43</td>
<td>- .03</td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td>.06</td>
<td>.26 *</td>
<td>.21</td>
<td>-.78 *</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td></td>
<td>-.64 *</td>
<td>-.14</td>
<td>- .13</td>
<td>-.15</td>
</tr>
<tr>
<td>Mod. Per.</td>
<td></td>
<td></td>
<td></td>
<td>-.64 *</td>
<td>-.13</td>
<td>-.13</td>
<td>-.04</td>
</tr>
<tr>
<td>Rpt. Per.</td>
<td></td>
<td></td>
<td></td>
<td>-.64 *</td>
<td>-.13</td>
<td>-.13</td>
<td>-.04</td>
</tr>
<tr>
<td>Biz. Resp.</td>
<td></td>
<td></td>
<td></td>
<td>-.64 *</td>
<td>-.13</td>
<td>-.13</td>
<td>-.04</td>
</tr>
</tbody>
</table>

PE= Total proportional error, r²= linear regression correlation squared, b= linear regression slope, Constant= linear regression constant, Mod. Per.= total modal perseverations, Rpt. Per.= total repetition perseverations, Biz. Resp.= total bizarre responses.

With a few exceptions the pattern of correlations is similar to that found in the whole combined group. Some of the correlations are similar but no longer significant (CET and PE, bizarre responses and constant, r² and modal perseveration). Two correlations were quite different. Bizarre responses were not significantly correlated with PE and r² as they were in the whole combined group. Some of the correlations in this group were of the opposite polarity to that in the whole combined group (CET and bizarre responses, CET and b, PE and repetition perseverations, PE and modal perseverations, r² and b, r² and repetition perseverations). The only one of these that was significant was the correlation between CET and b in the whole combined group. The strength of one correlation was greater and reached significance (PE and modal perseverations).

The correlations in the Non-BI group (n=92) are shown in table 14. Again, with a few exceptions the pattern of correlations is similar to that found in the whole combined group. Some of the correlations are similar but no longer significant (CET and PE, CET and r², CET and b, PE and constant, r² and constant, bizarre responses and constant). Several correlations in this group were of the opposite polarity to that in the whole combined group (CET and modal perseverations, CET and repetition perseverations, bizarre and modal perseveration, r² and b) although none of these correlations were significant.

The strength of two of the correlations was greater and reached significance (CET and bizarre responses, PE and $b$).

Table 14

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>.20</td>
<td>-.19</td>
<td>-.79</td>
<td>.13</td>
<td>-.12</td>
<td>-.70</td>
<td>.25*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.07</td>
<td>.18</td>
<td>-.12</td>
<td>.53*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.05</td>
<td>.09</td>
<td>-.23*</td>
<td>-.36*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.12</td>
<td>.02</td>
<td>-.10</td>
<td>-.60*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.13</td>
<td>.06</td>
<td>.18</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.01</td>
<td>.01</td>
<td>.10</td>
<td>.04</td>
</tr>
</tbody>
</table>

$PE$ = Total proportional error, $r^2$ = linear regression correlation squared, $b$ = linear regression slope, Intercept = linear regression constant, Mod. Per. = total modal perseverations, Rpt. Per. = total repetition perseverations, Biz. Resp. = total bizarre responses.

( * p<.05  † p<.01  ‡ p<.001)

There are several differences between the BI and Non-BI groups' patterns of correlations. Several of the correlations were not significant in the BI group that were significant in the Non-BI group (CET and bizarre responses, PE and $b$, PE and bizarre responses, $r^2$ and modal perseverations). Similarly, several of the correlations were not significant in the Non-BI group that were significant in the BI group (CET and $r^2$, CET and $b$, CET and constant, PE and constant, $r^2$ and constant). There were also several correlations whose polarity was opposite between the groups (PE and both perseveration measures, $r^2$ and repetition perseverations, bizarre responses and modal perseverations) although none of these correlations were significant.

Summary: The CET correlated significantly with the Milest measures of PE and $r^2$ in the whole combined group but not separate BI or Non-BI groups. Bizarre responding only correlated with the CET in the Non-BI group. In the BI group bizarre responding only correlated significantly with $b$. The perseveration measures correlated strongly with each other but generally were not correlated with the other measures with the exception of modal perseverations in the BI group and $r^2$ in the whole group and Non-BI group.

DISCUSSION

Test of Predictions

1. The brain injured participants will be worse at estimating which will be reflected in a significantly higher total proportional error score than the Non-BI groups on the first occasion.

The results support the first prediction that people with a brain injury will be less accurate in estimation. The BI group’s mean total PE score was significantly higher than the Non-BI group’s. Therefore the BI participants were worse at estimating the distances.

2. The BI participants’ poorer accuracy will also be reflected in the mean value of the linear regression constant which will be significantly larger than the Non-BI group’s.

The second prediction was not supported. There was no significant difference between the BI and Non-BI groups’ in the mean value of the constant.

3. The BI group will also be less consistent in their estimation performance and therefore their mean values of $r^2$ and $b$ will be significantly lower than the Non-BI group’s.

There was support for the third prediction. The BI group’s mean $r^2$ was significantly lower and mean $b$ was significantly higher than the Non-BI group’s. Because greater consistency would be indicated by values of $r^2$ and $b$ approaching one, this finding indicates that the BI group’s performance was less consistent.
4. The brain injured participants will benefit less from the supplementary information hence they will improve their accuracy of estimation less than the non-brain injured participants on the second occasion. However, if the brain injured participant’s poorer estimation performance is due to impaired access to semantic memory then the opposite prediction would be made.

There is not support for the prediction that the BI participants would benefit less from supplementary information. This would have required a significant three-way interaction between Brain Status, Occasion and Group. In other words those with a brain injury would have had to show a greater difference between the first and second occasion than the difference found in the Non-BI group. In order to rule out the possibility of a practice effect this improvement would need to have been found in the Experimental BI group but not the Control BI group who received no additional information.

In fact it was found that the BI group did improve but to the same extent as the Non-BI group. This improvement was only found in the Experimental groups, thus supporting the conclusion that it was the availability of the supplementary information rather than practice that underlies this effect.

The results do not support the semantic memory deficit hypothesis. The Experimental BI group did not improve to a greater extent than the Non-BI group on the second occasion.

5. There will be evidence of dysexecutive behaviour in the brain injured participants and therefore they will have higher scores on the CET and perseveration measures and will make more bizarre responses.

The fifth prediction was supported on the measures of total CET and number of bizarre responses. The BI group had significantly higher scores on both measures which is indicative of dysexecutive behaviour. However, there was no
significant difference between the BI and Non-BI groups on either measure of perseveration.

6. The supplementary information should have less impact on the Experimental BI group’s level of dysexecutive behaviour than on the Non-BI group’s.

There was support for the sixth prediction that the BI group’s dysexecutive behaviour would not benefit from supplementary information to the extent that the Non-BI group was able to benefit. However, the pattern of results is not exactly as predicted. The only three-way interaction that was significant was that for modal perseveration. There was not a significant interaction between Brain Status, Occasion and Group for repetition perseveration or bizarre responses. The Non-BI participants who received supplementary information did improve on the measures of modal perseveration and bizarre responding. In contrast the BI participants who received supplementary information did not improve on the these measures. Neither of these groups improved on the measure of repetitive perseveration.

7. There should be a significant positive correlation between the CET and the measure of bizarre responses on the Milest Test on the first occasion.

Contrary to prediction there was not a significant correlation between total CET and total bizarre responses.

8. Exploratory analyses; correlations.

The main Milest measures were significantly correlated with the CET score. There were differences between the BI and Non-BI groups’ patterns of correlation. Of particular interest were the finding that bizarre responding correlated with the CET in the BI but not the Non-BI group. PE did not significantly correlate with bizarre responding in the BI group but did so in the Non-BI group. The perseverative measures were not correlated significantly
with bizarre responding. The only Milest measure which was significantly associated with CET was $r^2$.

Critique of the Study: Complications and Limitations

Before moving on to discuss the findings in relation to models some potential limitations of the study should be considered. The analyses and results will considered after the measures and the sample.

The Milest does have advantages over the CET as a measure of estimation. Because the items all relate to the same domain it is easier to compare inter-item responses. This is unlike the CET which includes estimation of weight and height as well as non-numerical, non-interval responses (e.g. best paid job). The method of determining “extremeness” of response also has advantages. It is rational compared to Shallice & Evau’s (1978) method of defining the degree of extremeness. It is however, possible that the Milest Test’s power as a measure might be limited by the fact that the items are not evenly distributed across the range of distances (see figure 10).

![Figure 10](image_url)

Distribution of the Magnitude of Milest Test Items

There may also be a complication in the Milest Test methodology of providing supplementary information. In the Introduction it was suggested that providing supplementary information for the CET was likely to activate the Supervisory...
System. The separate sets of information needed for each item would be more salient than a single set of information and therefore more likely to trigger the System (see Shallice & Burgess 1991b). It was assumed that a single set of information, as is used in the Milest Test, would be less salient and therefore unlikely to activate the Supervisory System for more than perhaps the initial items. This assumption however may be false. It is possible that the supplementary information provided in the Milest Test is salient enough to activate supervisory functions. It might be argued that this accounts for the improvement on the second occasion in the Experimental BI group. If so then this method would not allow the semantic memory deficit hypothesis to be assessed. Improvement on the second occasion in the BI group might be due to the supplementary information compensating for executive impairment by activating the Supervisory System or it could be due to the information compensating for a deficit in access to semantic memories.

However, there is reason to believe that this has not happened. If the information did activate the Supervisory System then we would expect a reduction in the level of bizarre responses. This was not found. The Experimental BI group did not significantly reduce its level of bizarre responses on the second occasion. Therefore, assuming that bizarre responses are indicative of a Supervisory System failure when use of relevant information would prevent these, then it seems unlikely that the information was so salient that it activated the System.

The assumption implicit in the measurement of repetitive perseverations could weaken the reliability of this measure. It is assumed that the response used for one item is not suitable for the next item. However, it is possible that one estimate might suit two consecutive items if they were actually similar distances. For example, if two consecutive items were within 10 miles of one another (e.g. 56 miles and 65 miles) then one response could be an acceptable estimate for both items (e.g. “60 miles”). If this occurred frequently then the
incidence of repetitive perseveration might be over estimated. Fortunately there are only four instances where two consecutive items are within 10 miles of one another. However, the same response might be acceptable for even larger differences between two consecutive items. For example if the difference was 20 miles (e.g. 40 miles and 60 miles) then one response (e.g. “50 miles”) could be an acceptable estimate. There are in fact 10 pairs of consecutive items where the difference between the actual distances is 20 miles or less.

This factor could also affect the measure of modal perseveration because a single response might be accurate within ±10 miles for numerous items. Consequently repeated use of a particular response within the total set of items might in fact be an acceptable, approximately accurate response. In the Milest there are at least 16 items where the distances are within 10 miles of one another and there are at least 31 items where the distances are within 20 miles of each other. This situation complicates the interpretation of modal and repetition perseveration. Although they can not be interpreted as perseverations in the usual sense (e.g. in the Wisconsin Card Sorting Test perseveration is repetition of an previous error) it does not prevent them from being interpreted as meaningful behaviour which distinguishes the different groups.

There is also a potential weakness in the method of determining bizarre responses on the second occasion. On the second occasion the criteria for bizarreness was re-calculated based on the Experimental Non-BI group’s errors. Because they improved their accuracy on the second occasion it is possible that the criteria became more stringent. Consequently if the Experimental BI group actually reduced the magnitude of their errors the more stringent criteria might mask this effect.

The criteria of bizarreness in fact became more stringent on the second occasion for all the Milest items. If the criteria from the first occasion were applied to the second occasion then 59% of the BI group would be considered

to have improved (9% made the same number, 9% made more and 23% did not make any on the first or second occasion). However, 83% of the BI group who were considered to have made bizarre responses using the more stringent criteria were also considered to have made bizarre responses using the less stringent criteria. Therefore there are two reasons to consider this issue insignificant. Firstly, the difference in criteria would affect a minority of the BI group (i.e. 17% who might be considered not to have made any bizarre responses under the less stringent criteria). Secondly, and more importantly, the use of the more stringent criteria is justified as the point in question is whether they make more extreme errors compared to the Experimental Non-BI group’s performance on the second occasion.

Differences in sample sizes and characteristics of the groups (e.g. demographic differences) is another potential complication. Unequal sample sizes (i.e. the Control BI group containing 28 people compared to 46 in the other groups\(^{10}\)) can affect statistical analyses. Fortunately unequal sample sizes are unlikely to be problematic for repeated measures ANOVA as the cell sizes remain proportional (Howells 1997, p. 493). In the case of the correlations, there is a risk of Type II error (e.g. a weak correlation being rejected as insignificant) with smaller sample sizes.

Several results that suggest that there could be differences between the Experimental and Control BI groups in the nature and/or severity of their cognitive impairments. The Control BI group had a significantly higher level of modal perseveration on the first occasion. They also made significantly more bizarre responses on both the first and second occasions. The reasons for such differences might be clearer had additional measures of cognitive impairment been used with the two BI groups.

\(^{10}\) Ideally more participants would be recruited to equalise the group sizes, however, this was not possible in the time available. Data will continue to be collected to increase the sample size of this group.

The groups also differ in terms of demographic characteristics. However this can be addressed using statistical methods as was done here (i.e. using these variables as covariates). Since the observed effects were unchanged when the covariates were entered, we can conclude that these differences did not explain the between-groups differences.

Implications for the Supervisory Model Account of CET Performance

This study is particularly aimed at exploring the Supervisory System model account of the CET and the proposal that abnormal estimation performance gives us an indication of how the System can dysfunction. As described in the Introduction, the Supervisory Model is intended to account for peoples’ ability to respond appropriately to unique or unusual situations that can not be resolved by routine application of learned routines or schemata. An estimation task, such as the CET, is said to require the Supervisory System because it can not be solved simply by using “rote knowledge or a routine method” (Shallice & Burgess 1991a). The System is said to select or plan a strategy, evaluate its efficacy and make necessary modifications if the result is not adequate.

One suggestion is that people with frontal lobe damage perform abnormally on the CET due to impaired strategy selection. The Milest Test differs from the CET in that no more than one strategy is needed to solve all the items. The CET arguably requires more than one strategy because its items include different dimensions (e.g. height, age, weight and size) as well as different areas of knowledge (e.g. people, animals and buildings). Strategy selection may also be particularly important on the CET because the items are not grouped in a way that would allow a strategy to be selected and used with several items in sequence. Hence selection of a strategy would need to occur several times during the CET. However, a single strategy can be applied to all the Milest items and therefore once it is selected further strategy selection is not necessary. Considering this the strategy selection hypothesis would predict that

people with frontal dysfunction would be impaired on the CET yet perform normally on the Milest. However, the results do not support this. The BI group is worse than the Non-BI group on both the CET and the Milest Test. The fact that the BI group can produce relatively accurate estimations on some, if not many of the Milest items is indicated by relatively low mean bizarre response score (i.e. less than 10% of the total items). Therefore, they must be capable of selecting an appropriate strategy. Consequently, the strategy-selection impairment account is not supported by the results of this study. However, the results do not preclude the possibility that the BI group is less able to maintain an appropriate strategy in order to apply it consistently.

Another suggestion is that frontal lobe damage impairs Supervisory evaluation and modification of plans or actions, resulting in abnormal CET performance. If bizarre estimations are considered to be instances of such Supervisory failure then the BI group’s performance on the Milest supports this account. They produce significantly more bizarre responses than the Non-BI group. Moreover, the Experimental BI group continues to produce more bizarre responses than the Non-BI group on the second occasion when the supplementary information makes it possible to avoid making bizarre responses. The potential of this information to achieve this is indicated by the fact that the Experimental Non-BI group reduce their mean bizarre estimation score to about one. Therefore the results support the evaluation/regulation impairment explanation.

The prediction made on the basis of the semantic memory deficit account of CET performance was that the Experimental BI group would improve to a greater extent than the Non-BI group on the second occasion. This was based on the argument that the supplementary information would essentially place the two groups on an equal footing because the BI group would not need to access semantic memory for relevant information. However, contrary to prediction it was found that the BI group did not improve to a greater degree. As the BI
group's performance is still significantly worse than the Non-BI group's we can conclude that the difference is not due to impaired access to semantic memory.

This interpretation is weakened by the multi-faceted nature of estimation tasks and the heterogeneity of the BI group. As mentioned in the *Introduction*, the fact that multiple factors were found in factor analysis of the CET indicate that it involves several cognitive abilities. There is no strong reason to believe that the Milest Test is cognitively pure. Logically we would expect an estimation task to be multi-faceted, involving various cognitive systems and skills such as semantic knowledge, numerical information, executive skills (see below). although the BI group was selected to have impairment of frontal, executive functions this was not necessarily their only impairment. Causes of brain injury such as closed head injury and anoxia are likely to affect several cognitive functions. Because these were not controlled (e.g. use of other measures of impairment as covariates) we can not sure of the extent to which they contributed. Therefore it is possible that the reason the Experimental BI group did not improve more than the Non-BI group because of other cognitive impairments. In fact the semantic memory hypothesis suggests that impaired CET performance is due to a combination of impairments (Taylor & O'Carroll 1995).

*Developing a Theoretical Model*

As mentioned in the *Introduction*, estimation tasks have been discussed in relation to models of how judgements are made in the reasoning and problem solving process. There are two broad areas of literature that are relevant; that concerning the Supervisory System and how it is affected by frontal damage and the literature concerning judgement processes in non-injured people. As described earlier, the Supervisory System was proposed to explain how people are able to respond to non-routine situations requiring a novel solution.

There is a considerable volume of literature concerning normal judgement processes. Some of this literature relating to how people estimate things was summarised by Brown & Seigler (1993). One proposal is that people use "logical shortcuts such as rules of thumb" (Mayer 1992), referred to as heuristics. Various heuristics have been described including the availability heuristic. Use of this heuristic would lead the person to base their estimate on the information that was most available to them (i.e. that which is most easily recalled). What have been referred to here as perseverative responses might arguably arise from use of the availability heuristic. Assuming that the most recent response is still activated, it could be selected because it is the most available response. It would make sense to assume that this would occur (or is more likely to occur) when no other potential response has been activated to a greater degree.

It has also been proposed that people use domain-specific knowledge upon which to base an inference. In this study the experimental groups had such information available to them. It is safe to assume that this would therefore be the best strategy and hence the most likely one to be used on the second occasion (unless they could recall the exact distance). If this was the case it is interesting that people (both BI and Non-BI groups) continued to make bizarre as well as perseverative responses. This suggests that they did not consistently apply this strategy to all the items. In the case of repetitive perseverative responses this would be difficult to argue in view of the complication with this measure mentioned above. However, a bizarre response would clearly be an instance in which there had been failure to use the domain-specific knowledge. This suggests that once a strategy is selected it does not automatically remain 'switched on' until 'switched off'. On the second occasion when a bizarre response occurs following several approximately accurate responses it must have occurred because the previously successful strategy has not been reapplied. Even though such a response indicates a failure of response monitoring and regulation, it also implies inappropriate strategy application as

the very inaccurate response would not have been selected had the previously used strategy been re-applied.

In view of this we could postulate that strategy selection must occur for each instance, with perhaps a bias to select the most recent successful strategy. However, this would be quite uneconomical. Variation seems more likely to occur because situations are not stable (e.g. other distractions occur) and people do not behave in a mechanical manner. Therefore such biasing of strategy selection to re-apply the recent strategy could occasionally be disrupted, thus resulting in intermittent errors. We can extend this proposal to include some sort of supervisory process capable of monitoring the estimation process in order to trigger some rectification when this occurs. If as Shallice, Burgess and others have argued, dysfunction of the frontal lobe activities causes impairment of supervisory processes, then such errors would be more likely in such individuals. The results of this study support such an account. This account could also explain the behaviour of some BI patients on the Wisconsin Card Sorting Test. ‘Failure to maintain set’ has been described and found to be more common in people with frontal lobe damage (Heaton 1981). It refers to errors following a series of correct responses. Therefore it too is a case where the application of an appropriate strategy has been interrupted.

Before going on to describe a model that might account for the findings, it is worth summarising the main points. The non-BI group is inaccurate when estimating inter-city distance but improve their accuracy when supplementary information is available. However, they continue to be quite inaccurate. The BI group are more inaccurate at estimating distance than the Non-BI group. Contrary to prediction, they are able to benefit to the same degree as the non-BI group from the supplementary information. Although they reduced their total error score they continue to be worse than the Non-BI group. The pattern of responses in the BI group is also different from that in the Non-BI group. They did not significantly reduce the number of ‘extreme’ errors (i.e. bizarre
responses) that they make and they more often repeat a previous response, either using the same response in succession or using it repeatedly throughout the task.

Although the main evidence is from the proportional error scores (PE), the data from the linear regression allows further exploration of estimation performance. The three regression measures are inter-related but reflect different facets of accuracy. On the first occasion the BI group is worse on all four measures with the exception of the linear regression constant (i.e. PE, $r^2$ and $b$). They tended to over-estimate to the same degree as the Non-BI group. In the Experimental BI group the benefit of the supplementary information was reflected only in the measures of PE and constant. They did not improve on the measures of $r^2$ and slope which reflect consistency of responding. In other words, although they were able to use the information to enhance their accuracy, they continued to be less consistent than the Non-BI group.

The bizarre response measure is worthy of special consideration. Although it has been defined here as a measure of executive dysfunction it overlaps with several of the measures of accuracy. It seems appropriate, in retrospect, to consider bizarre responses as a measure of accuracy as by definition they are errors, albeit ones of unusual magnitude. They are therefore simply errors that have been singled out due to their degree. Hence when we try to account for them we need to consider that they are essentially errors as well as the fact that they are unusually large errors.

We would therefore expect bizarre responses to affect total PE. Although they do correlate in the Non-BI group ($r=.53$), in the BI group they do not correlate ($r= -.03$). One possible account for this pattern of results is that when there is no supplementary information bizarre responses can reflect different phenomena. In the Non-BI group they, like all the other errors, arise from lack of relevant knowledge. Hence PE and bizarre responding correlate. In the BI

group they occur during the first occasion not only because of lack of knowledge but also because of executive impairment.

As discussed above, it is a misnomer to refer to some responses as perseverative. They are indicators of the pattern of response behaviour. However, they are not necessarily a repetition of a previous error, which is the common definition of perseveration. What these ‘perseverations’ reflect is a relative lack of variety or narrowness of responses. For example, if someone gave a unique and different response for each item there would be no modal response and no repetitious responses. They would be using a wide and varied response set. However, the more modal responses and/or the more repetitious responses there are, the narrower and less varied the response set.

The model that aims to encompass these results and current models can be described as follows (see figure 11). An estimation question activates processes (in parallel) that result in the specification of the features of the number that would broadly suit the task (e.g. range, randomness, etc.). Based on subjective experience, these processes often occur very rapidly and automatically. This leads to generation of a number based on the defined features. This process is likely to be an iterative one of ‘successive approximation/estimation’. In this either the initial potential response or possibly several competing responses are compared with the specification to assess their adequacy.

When a satisfactory number is identified it is activated as the potential response for output. This is represented in the model as the ‘Pre-Response Buffer’. This is a temporary state (i.e. it occurs just prior to the overt response) and is not necessarily meant to indicate a type of memory that holds the response before it is actioned. However, such a component might be involved in circumstances where it is used to vet the response (e.g. when there is a high risk or cost contingent on errors, such as oral examinations or game-shows).

The model also depicts two other factors that can impinge on the process. One is the influence of a supervisory process. This includes the type of biasing mechanism described by Shallice et al. This would prevent the output resulting from an un-checked routine or habitual response. This biasing could act upon the strategy selection. For example, perseverative behaviour would occur if the availability heuristic was activated automatically leading to the most recent response being selected. It could also bias the iterative number generation-comparison process. Biasing might result either in low or high completion
threshold. For example, a slow, deliberate, careful approach would result from biasing of a high selection threshold. Alternatively, a fast or impulsive approach would result from biasing of a low threshold. Biasing processes might also affect the pre-response stage. This would enable the potential response to be monitored and regulated if necessary.

The other impinging factor that has been included is an evaluative or 'emotional' influence. On a general level it is well established that evaluative processing of information can result in behaviour being adversely affected by emotions (e.g. performance anxiety). They are also mentioned here because subjective feedback suggests that selection of an acceptable response involves a "feels right" judgement. The source of this is unknown. However it could be speculated that it originates in part from unconscious / pre-conscious activation of relevant knowledge. For example, knowledge that was activated by the question or by other influences (e.g. task constraints or rules) could influence the process in a way that was "felt" rather than "known". The influence of such 'intuitive' processes has also been proposed by Claxton (1998).

The impairment of such processes has been identified in people with frontal lobe damage. They display a dissociation between knowledge and action resulting in 'breaking the rules'. Although they know what they should or should not do, their behaviour is not guided by this. They do not display emotional conflict or distress that might normally arise from this contradiction. In other words the selection of responses is not affected by "feels right" (or "feels wrong") influences. It has been suggested that frontal lobe damage tends to affect the heteromodal cortices which integrate information from various sources and that consequently this damage produces typical 'frontal phenomena' (Prigatano 1991). Such phenomena would include impaired insight as well as 'bizarre' estimates.

The fact that errors are a ‘normal’ occurrence in human reasoning and problem solving is well known and extensively studied (see Brown & Seigler 1993). It has been argued that the level of error resulting from ‘fast and frugal’ judgement is not only sufficient but often surprisingly accurate (Gigerenzer et al 1996; Chater et al. in preparation). On the first occasion the degree of estimation error is likely to depend largely on more practical constraints, such as knowledge limitations. However, the persistence of error on the second occasion when supplementary information is available is more likely to arise from the so called ‘fast and frugal’ tendency of human judgement processes.

An important question is why the BI group is more inaccurate than the Non-BI group. On the first occasion it seems unlikely that such error arises from deficient specification of the broad features for the response. They do not tend to make improbable estimates. On mainland Britain no distance is likely to exceed 600 miles and none will exceed 800 miles. That the BI group does not make such errors indicates that they have specified a reasonable domain. Hence they are at least applying ‘intuitive statistics’ (Brown & Seigler 1993).

Unfortunately the results from the first occasion do not allow us to infer the possible source of the supra-normal error in the BI group. We can speculate that it arises from inadequate application of strategic processes or semantic knowledge. This could be due to impaired supervisory biasing of these processes. For example, failure to de-select inappropriate knowledge or heuristics and/or failure to select appropriate knowledge or heuristics. It is also possible that error arises from impaired supervisory regulation of the iterative process of response selection. This could result in premature, impulsive selection and output of a response.

The fact that the Experimental BI group improves to the same extent as the Experimental Non-BI group suggests that they are able to select an appropriate strategy. In other words they apparently are able to use the most pragmatic

strategy of using the available information to aid inductive reasoning. Because they continue to be worse than the Experimental Non-BI group indicates that their impairment is not simply due to deficient access to semantic memory. However, the results suggest that they are less consistent in the application of a strategy. They continue to display a more narrow, less varied pattern of responses, when the available information should allow a more varied set of responses. Indeed, the Experimental Non-BI group reduces the degree of ‘modal perseverations’, thus indicating a more varied response set. The fact that they do not reduce the number of ‘repetitive perseverations’ is not problematic to this account in view of the discussion above about the possibility that these are more likely to reflect two items that could reasonably be answered with the same estimation. Certainly the low number such responses (mean score = 2.09) supports this.

An equally credible explanation of the lower accuracy of the Experimental BI group on the second occasion is that their regulation of the response selection and monitoring is impaired.

Finally, it is worth considering the clinical sensitivity of the CET and Milest Test. Several weaknesses affecting the CET’s sensitivity were discussed in the Introduction (e.g. low split half reliability, multi-factorial structure). The Milest Test measures were not found to be better than the CET at distinguishing ‘good/bad’ performance with the exception of the squared regression correlation ($r^2$). The measure of bizarre responses might be clinically useful considering that the BI participants were significantly different from those without a brain injury on this measure. However, the means of the two groups were low and very similar (i.e. 3 versus 2; BI and Non-BI). Thus this measure has the problem of a low ‘floor’ as well as closely overlapping distributions that would make it clinically insensitive.

SUMMARY

People with a brain injury affecting the frontal lobe executive functions were found to be impaired on two estimation tasks, the CET and the Milest Test. They were able to benefit from the provision of information and so improve the accuracy of their Milest estimations. However, they were still less accurate than people without a brain injury. The results do not support the suggestion that impaired access to semantic memory underlies impaired CET performance. They do however support the proposal that abnormal CET performance is due to impairment of supervisory processes. Impaired strategy selection is only likely to be one contributing factor because even when strategy selection was not necessary (i.e. when supplementary information was available) people with a brain injury were still less accurate. It is proposed that the difference between the BI and Non-BI groups is that the former behave less consistently as a result of impaired supervisory processes.

References


Halstead, WC (1947) *Brain and Intelligence*. Chicago: Chicago University Press.


Wechsler, D (1981) *The Revised Wechsler Adult Intelligence Scale*
### Cognitive Estimation Test (CET)

<table>
<thead>
<tr>
<th>Name:</th>
<th>Date:</th>
<th>Examiner:</th>
<th>Place:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(NB The signs &gt;; &lt; are to be interpreted strictly i.e. not as ≥ and ≤)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 What is the height of the British Telecom Tower?</td>
<td>&gt; 1500'</td>
<td>= 1500'</td>
<td>&gt; 800'</td>
<td>100 - 800 ft.</td>
<td></td>
</tr>
<tr>
<td>2 How fast do race horses gallop?</td>
<td>&gt; 50 mph</td>
<td>= 50mph</td>
<td>&gt; 40 mph</td>
<td>&lt; 40 mph</td>
<td></td>
</tr>
<tr>
<td>3 What is the best paid job or occupation in Britain today?</td>
<td>manual workers</td>
<td>car workers (or well paid blue collar worker)</td>
<td>professional (to &amp; including PM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 What is the age of the oldest person in Britain today?</td>
<td>&gt; 115</td>
<td>= 115</td>
<td>= 114</td>
<td>104 - 113</td>
<td></td>
</tr>
<tr>
<td>5 What is the length of an average man's spine?</td>
<td>&gt; 5'</td>
<td>= 4'</td>
<td>= 4'</td>
<td>&gt; 1' 6'' - &lt; 4 ft.</td>
<td></td>
</tr>
<tr>
<td>6 How tall is the average English woman?</td>
<td>&gt; 5' 2''</td>
<td>= 5' 11'' = 6'</td>
<td>= 5' 9'' = 5' 10''</td>
<td>5' 3'' - 5' 8''</td>
<td></td>
</tr>
<tr>
<td>7 What is the population of Britain?</td>
<td>&gt; 1,000 million</td>
<td>&gt; 500 million</td>
<td>= 500 million</td>
<td>11-499 million</td>
<td></td>
</tr>
<tr>
<td>8 How heavy is a full pint bottle of milk?</td>
<td>&gt; 3 lbs.</td>
<td>&lt; 3 lbs.</td>
<td>= 3 lbs.</td>
<td>1 lbs. 1 oz. to</td>
<td></td>
</tr>
<tr>
<td>9 What is the largest object normally found in a house?</td>
<td>&lt; carpet</td>
<td>carpet</td>
<td>piano, settee, sideboard</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 How many camels are there in Holland?</td>
<td>Very large number</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Norms (O’Carroll et al (1994))** 5.3 (± 3.6)  
**CET Total**  
<table>
<thead>
<tr>
<th>&gt; 13</th>
<th>&gt; 9</th>
<th>2 - 9</th>
<th>&lt; 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; − 2 sd</td>
<td>&gt; − 1 sd</td>
<td>Average Range</td>
<td>&gt; ± 1 sd</td>
</tr>
<tr>
<td>&lt; 5th centile</td>
<td>&lt; 15th centile</td>
<td>15th – 85th centile</td>
<td>&gt; 85th centile</td>
</tr>
</tbody>
</table>
READ THIS BEFORE LOOKING AT THE REMAINING PAGES

There are 3 parts to this questionnaire. It is important that you do each part separately, without looking forward to the other parts or looking back at the parts you have already done. Please do not refer to any other books or maps while answering the questions.

All 3 parts must be done in a single period without any break or interruption.

Open to the next page.
INTER-CITY DISTANCES (1)

Below you will see a list of pairs of cities. Please write what you think the distance is between the cities (in miles) travelling by road. Although you may know the actual mileage for one or two, you will probably have to guess or estimate the distance. Please do all of them and when you reach the end check that you have not left any blank. Thank you.

London - Brighton _____  Oxford - Cardiff _____
Plymouth - Penzance _____  London - Birmingham _____
Birmingham - Oxford _____  Newcastle - York _____
Preston - Liverpool _____  Oxford - Nottingham _____
Manchester - Leeds _____  Edinburgh - Newcastle _____
Newcastle - London _____  Birmingham - Manchester _____
Birmingham - Liverpool _____  Plymouth - Southampton _____
Manchester - Newcastle _____  Manchester - Oxford _____
Plymouth - Exeter _____  London - Glasgow _____
Preston - Manchester _____  Newcastle - Southampton _____
London - Oxford _____  Preston - Sheffield _____
Manchester - Liverpool _____  Carlisle - Newcastle _____
Bristol - Plymouth _____  Edinburgh - Aberdeen _____
Birmingham - Sheffield _____  Preston - Birmingham _____
London - Cardiff _____  Aberdeen - Fort William _____
Edinburgh - Glasgow _____  London - Edinburgh _____
Oxford - Bristol _____  Preston - Carlisle _____
Birmingham - Newcastle _____  Newcastle - Leeds _____
Manchester - Sheffield _____  Cardiff - Bristol _____
Plymouth - London _____  Edinburgh - Fort William _____

Please check that you've not missed any and then turn the page over so this is out of sight.
Leeds - Sheffield = 36 miles
London - Cambridge = 60 miles
Bristol - Edinburgh = 381 miles
INTER-CITY DISTANCES (2)

Now I would like you to use the map to help you estimate the distances (Note: These are the same pairs of cities you saw previously. Remember, you should not look at the distances you wrote on the previous sheet). Thank you.

London - Brighton
Plymouth - Penzance
Birmingham - Oxford
Preston - Liverpool
Manchester - Leeds
Newcastle - London
Birmingham - Liverpool
Manchester - Newcastle
Plymouth - Exeter
Preston - Manchester
London - Oxford
Manchester - Liverpool
Bristol - Plymouth
Birmingham - Sheffield
London - Cardiff
Edinburgh - Glasgow
Oxford - Bristol
Birmingham - Newcastle
Manchester - Sheffield
Plymouth - London
Oxford - Cardiff
London - Birmingham
Newcastle - York
Oxford - Nottingham
Edinburgh - Newcastle
Birmingham - Manchester
Plymouth - Southampton
Manchester - Oxford
London - Glasgow
Newcastle - Southampton
Preston - Sheffield
Carlisle - Newcastle
Edinburgh - Aberdeen
Preston - Birmingham
Aberdeen - Fort William
London - Edinburgh
Oxford - Carlisle
Newcastle - Leeds
Cardiff - Bristol
Edinburgh - Fort William

When you have finished, turn the page so this is out of sight.
In order to understand the results of these questionnaires I also need to gather the following information. This will help to explore the differences between people in their ability to know or guess distances. All the information will be kept confidential. If you prefer not to give your name please leave the space blank.

Name ___________________________ Date ______________________

City (on the map) nearest to your house ________________________

Circle the answer which applies to you:

I am a car driver: Yes / No

I am in the following age group:

16-25  26-35  36-45  46-55
56-65  66-75  75+

My formal education finished with:

No exams done  GCSE/GCE
0-levels  University Degree
A-levels  Postgraduate Degree
National Certificate  Other Qualifications

For Official Use:

Ctr / BI (CVA / CHI / O)

CET Total _______ CET Item Scores ______________________ NART _______
The Contribution of Image Change

and Role Fulfilment in the Present and Past

to the Psychological Well-being of Spouses

Caring for their Dementing Partner

Drew Alcott
Exeter District Department of Clinical Psychology
April 1986
Research Dissertation submitted in partial completion
of the MSc. in Clinical Psychology required by
the University of Exeter, Devon.
### TABLE OF CONTENTS

Acknowledgements  
Abstract  
Introduction  
Literature Review  
- The Consequences of Care-giving  
- The Determinants of Strain in Carers  
- Demands involved in direct care  
- Intrapersonal demands of care-giving  
- Interpersonal demands of care-giving  
- Demands concerning interaction with social and health care networks  
Mediators of strain  
Linking Section  
Design  
Methods  
- Informants  
  - Carers  
  - Controls  
Measures  
- Repertory grid  
- GHQ  
- Strain scale  
- Behaviour & mood disturbance scale  
- Ischemia scale  
- CAPE  
Procedure  
- The dementia group  
- The control group  
Results  
- Descriptive data  
- Change of Image(IC)  
- Role fulfilment before(RFB)  
- Current role fulfilment(RFN)  
- Strain  
- GHQ  
- Esteem  
- Behaviour & Mood Disturbance  
- Other analyses  
Discussion  
- Image change  
- Current role fulfilment  
- Previous role fulfilment  
- Measures of psychological well-being  
- Particularly important constructs  
- Clinical implications  
Conclusion  
References  
Appendices
ACKNOWLEDGEMENTS

There are a number of people whose co-operation made the completion of this project possible. I wish to repeat my thanks to all the carers who were willing to spend their free time with me. I am also grateful to the older people in the various housing schemes and groups, particularly the members of the Whipton Over-60's Club, who helped to make up my comparison group. Contacting carers would not have been possible without the help of David Jeffery, Julian Herbert, and other members of the Torbay Team for the Elderly as well as Dr. Gordon Langley and Mike Aitken at Exminster Hospital.

I would like to acknowledge the support and guidance I received from various people. The amount of time and energy which was consumed by this project could not have been spent without the tolerance and support of my wife, Gill. I am also grateful to Paul Webley whose reliable supervision gave me the sounding board and direction I wanted. I also appreciated the useful suggestions I got from Neil Matson who also helped to kindle my interest in the elderly. Rachel Kirby deserves thanks for her tireless contribution to helping me grapple with the computer. And finally I am grateful to the other trainees with whom I have shared the past 3 years for all the little bits of interest and encouragement which added to the total experience.
ABSTRACT

Previous research has studied various factors relating to the level of strain in those who are caring for a dementing relative. However relatively little of this has concerned psychological factors despite the fact that numerous of these have been suggested in the literature. The present study has used repertory grid technique to focus on several of these including the change in the carer’s Image of their dementing partner and the degree of role fulfilment by the partner both present and past.

The predicted relationship between Image change and strain was not found. Image change was associated with mental health overall but this was due to the implications of these changes for role fulfilment. It was found that spouses' Image of their partner does change and amongst carers the nature and extent of change was related to the sequelae of dementia. The degree of current role fulfilment was associated with mental health. Mental health was poorer when the degree of role fulfilment was low, regardless whether the person was caring for a dementing partner. Compared to others of a similar age, the degree of current role fulfilment was lower amongst those whose partner was dementing. Particularly important aspects of role fulfilment concern how caring, open, and reciprocative the partner is. The degree of previous role fulfilment was also related to carers' psychological well-being. In particular this concerned how caring, helpful, calm, patient, and able to remember things the partner had been. It is suggested that this reflects the quality of the premorbid relationship in terms of how positive and nurturative it was. Two other important constructs related to carers' psychological well-being were Contentedness and Goodness. Various possible reasons for this are suggested.

The results of this study have implications for those working with carers. These concern the assessment of their psychological well-being and the factors contributing to this as well as possible areas for counselling and support. The various measures used to assess psychological well-being were not found to be equivalent. This indicates the value of using more than one such measure when studying factors related to strain in carers.
INTRODUCTION

With improvements in social and medical practices there has been an increase in our average life expectancy. Consequently there has been an increase in the number of elderly people in the population.

Approximately 5-10% of the older population are believed to be suffering from an organic dementing process (Zarit et al 1980). The incidence in those over the age of 80 years old is even higher at around 22% (Kay et al 1964). The degree of severity varies with deterioration being mild in about 5.2% and severe in 4.9% of those 65 and older (Kay et al 1964). Clearly then there is a sizeable number of elderly people in the population suffering from dementia.

Yet the majority of this group are able to continue to live in the community. Kay et al (1964) estimate that 83-90% of the dementing elderly are cared for in the community. It does not seem that these represent the less severely demented cases as they found that those living in the community were as severely demented as those in hospital. As Gilleard et al (1984,b) state, "the community remains and will continue to remain the principal care setting", which they point out, is government policy.

Although the primary care-giver, whose efforts allow the dementing person to remain in the community may be any relative, the majority are spouses (Zarit 1982). Based on her review of the literature, Zarit estimates that 45-56% of all carers are spouses. The research reported here has addressed this group rather than all carers which would include those who are the "children" of the dementing person.
A great deal is involved in the process of caring for someone who suffers with dementia. The carer faces a number of demands and coping with these can have various consequences. These will be described in more detail later. It is sufficient at this point simply to note that they involve not only financial costs but also social, psychological, and physical costs. As a result, the carer often experiences some degree of strain.

Strain is one of several factors which influence carers' willingness to continue to provide the care needed to allow the dementing person to remain in the community. Maintaining the dementing person in the community is important for at least two reasons.

In economic terms it is crucial to the quality of institutional services. These are already stretched and could not cope with the flood of patients which would result if many carers decided they could no longer cope with the demands.

The dementing person's quality of life is also important. The effect of relocation (e.g. from community to hospital) can itself have negative consequences (Blenkner 1967, cited in Zarit et al 1980). In general it is accepted that dementing patients do best when maintained in their own home for as long as possible (Kennie & Moore 1980, cited in Zarit 1982). For these reasons it is important for professionals to know about the mediating factors of strain in carers.

A number of these mediating factors have been revealed by research which will be outlined later. These relate to the carer (e.g. their own health), the dementing person (e.g. the type or degree of behaviour problems), and others involved with the carer (e.g. support from family members).
However, in general it is the qualitative rather than the quantitative aspects of these which are most important. For example, it is not how often family members visit or how severe the dementia is which is crucial, but how the carer perceives these things.

This is the aspect of carers' experience which will be addressed here. The focus of this research has been the psychological variables which mediate between what is happening and how the carer reacts to this. In particular I have studied the carer's Image of their dementing partner and the extent to which they are perceived to fulfil their role at present as well as in the past.

LITERATURE REVIEW

THE CONSEQUENCES OF CARE-GIVING

Providing care for a dementing relative places numerous demands on the carer which may have various consequences. The findings of Sanford (1975) suggest that these fall into three categories.

As their own health may be affected physical consequences are one category. Sanford (1975) found that carers complained of arthritis, back strain, respiratory difficulty, and lacking strength for lifting. Although these were not necessarily caused by care-giving there is little doubt that the physical and emotional demands on carers will affect their health and may cause or exacerbate medical problems.

Care-giving may also have social consequences. Almost half of those in Sanford's study said that their social life had been restricted. This may be because they become tied to their home as their partner becomes increasingly dependent. Over a
quarter of Sanford's group said that they were no longer able to leave their relative alone for more than an hour. Several in his group also said that they had had to give up their job in order to care for their relative. As well as reducing their social contacts this sometimes had financial disadvantages.

The third category is that of psychological consequences. Depression, anxiety, and embarrassment were reported by some of the carers in Sanford's (1975) study. Psychological effects of care-giving may arise from personality conflicts between the carer and their relative. A quarter of Sanford's carers reported this as a problem. It is also likely that social and physical consequences will have psychological effects.

I will go into more detail about the psychological consequences of care-giving next and go on to focus on the general level of psychological strain which carers experience in the process of caring for their partner.

The Psychological Consequences of Care-giving

The psychological consequences of care-giving are mentioned throughout the literature, however an article by Barnes et al (1981) provides a good summary of these. They describe the sorts of problems which carers talked about while taking part in a carers support group. These refer to various aspects of the carers' experience.

The diagnostic process itself can have adverse psychological effects on carers. When they are told by the doctor what is wrong with their partner they can have difficulty taking in this information because of their shock or because it was not fully or clearly presented. This inadequate understanding may contribute to a sense
of confusion in the carer. Being told the diagnosis and prognosis may evoke feelings of hopelessness in them. It can also lead to a sense of uncertainty about the future as they realise that their plans and expectations have been shattered.

As the reality of the situation sinks in and they observe changes in their partner, carers often experience a sense of loss. Barnes et al (1981) conceptualise this as a process of 'chronic grief'. Carers have feelings similar to those common to acute grief such as sadness and depression. Although their partner is physically present and looks more or less like they always have they are mentally quite different and the carer feels the loss of the person they have known.

The dementing person's behaviour may also evoke strong feelings such as feeling embarrassed by their partner particularly when in public. They may become frustrated or angry when their partner is unable or appears unwilling to do things which previously they were able to do. Carers can feel quite resentful about what their partner does or does not do. It may also relate to things which they must now do or plans which have had to be abandoned.

As their partner deteriorates they will require more care and supervision. With increasing absent-mindedness the carer may become fearful or anxious about their partner's safety. When constant supervision is needed the carer may feel tied to them (Gilleard et al 1984,a). Dependency as well as embarrassment may result in the carer becoming socially isolated and lead to feelings of being trapped at home.

Feelings of guilt are not uncommon amongst carers. This may be because they feel responsible in some way for causing their partner's disturbed behaviour. It may also relate to a sense of inadequacy about their ability to help their partner whom they
see continuing to deteriorate despite their efforts. A sense of guilt may also relate to the feelings which they have about their partner or their circumstances. They may think that they should not feel frustrated, angry, or resentful because they are aware that their partner is not personally responsible for what happens.

It is clear from this that care-giving may have various psychological effects on carers including feelings of confusion, anger, frustration, resentment, anxiety, and loss.

Many of these relate to the changes taking place in their relative. Together with other consequences these contribute to the overall impact of care-giving on the carer. Essentially this concerns their ability and willingness to continue to maintain their partner in the community.

Strain and Burden in Carers

Studies of carers have tended to consider the overall impact of care-giving under rubrics such as 'burden', 'strain', 'morale', or 'distress'. Different measures have been used to reflect these general states.

Carers' level of morale was measured by Gilhooley (1984) using the Kutner Morale Scale which has been shown to correlate with other measures of adjustment and life satisfaction. She also used a 'mental health' scale from the OARS Multidimensional Functional Assessment Questionnaire which involves the interviewer rating the carer on the basis of their responses to questions from the Present State Examination Schedule (PSE). Together these were taken as indicators of the carer's 'psychological well-being'. Gilhooley found that carers' level of morale tended to be rather low although in general their mental health was good or only mildly 'impaired'.
Mental health measures have also been used to reflect the level of 'emotional distress' in carers. Gilhooly et al (1984, b) used the General Health Questionnaire (GHQ) for this purpose. Although this is a measure of (psychiatric) 'disturbance' they argue that it can be used as a measure of 'distress' when used to compare scores in different groups. Although they did not find significant differences between groups (e.g. carers whose relative was attending day hospital vs. those whose relative was not) they found overall that at least 50% of the carers scored above the 'caseness' threshold on the GHQ. This indicates that many of the carers were experiencing psychiatric disturbance if not distress.

Zarit (1982) also used a measure of mental health as an indicator of distress amongst carers. The Brief Symptom Inventory (BSI) which she used is a measure of psychological disturbance or 'psychopathology' consisting of 31 items from the Symptom Checklist (Derogatis et al 1970). Only the female carers' scores on the BSI were higher than those of a matched control sample from the general community.

Burden has been measured using a specially designed self-report inventory (Zarit et al 1980). Their 29 item measure consists of statements derived by Zarit et al from clinical experience. It includes various things such as carers' reactions to their partner's behaviour (e.g. "I feel embarrassed over my spouse's behaviour"), social consequences (e.g. "I feel my social life has suffered because of my involvement with my spouse"), and relationship conflicts (e.g. "I feel strained in my interactions with my spouse"). Zarit et al were surprised that 'the level of burden measured was less than expected'. Interpreting this is difficult because there is no comparison group.
Strain has been measured in different ways. Gilleard et al (1982) asked carers to rate each item of a problem checklist (concerning the dementing partner) in terms of its degree of burden for them and the degree to which they felt they could cope with it. They combined these responses to indicate the carer's level of strain.

Greene et al (1982) measured strain using a 15 item questionnaire which they designed for this purpose. Like Zarit et al's (1980) burden inventory, Greene's 'stress scale' included various statements which carers replied to by indicating frequency (5 point scale from "Never" to "All the Time") or severity ("Not at All" to Considerably") of items.

Their factor analysis of the results indicated that the items were clustered in three groups. These were 'personal distress' (e.g. "Do you ever feel you can no longer cope with the situation"), 'domestic upset' (e.g. "How much has your, social life been affected"), and 'negative feelings' (e.g. "Do you ever get cross and angry with...."). Again it is difficult to interpret their results as no comparison group was used. Although it is possible to compare one carer with another, it does not allow us to say how much strain they experience compared to non-carers.

Gilleard et al (1984,a) also used a specially designed strain scale. It consisted of 13 questions, each of which concerns a particular component of strain (e.g. depression, ability to cope, difficulty in sleeping, feelings of embarrassment, etc.). It is not clear what the results of this measure were other than how it related to other measures (e.g. a problem checklist).

The question which arises is whether these measures are equivalent and whether they reflect the same underlying phenomena. This can be considered by looking at those studies which used more than one of these measures.
In her study, Gilhooley (1984) used a measure of morale as well as a measure of mental health. On the whole the results of both measures were in the same direction i.e. showing lower morale and poorer mental health in relation to particular variables (e.g. the duration of care-giving). However there were some differences. For example, she found that the 'distance' of the blood relationship between the carer and the dementing relative correlated with mental health but not morale (the greater the distance the better the carer's mental health).

In her study of carers, Zarit (1982) used both a measure of mental health and a measure of burden. Although the results of these measures were correlated they did not always both correlate with other measures. For example, the carer's willingness to consider putting their relative in hospital was correlated with the degree of burden but not with the level of distress. Carers whose degree of burden was high were more likely to consider hospitalisation while those whose level of strain was high were not necessarily so inclined.

Gilleard et al (1984,a) used a measure of strain as well as measures of burden and distress. Unfortunately their analysis does not shed any light on the relationship between these three measures as they combined the results of the strain and burden measures and did not report the results of the measure of distress (GHQ) in this article (cf. Gilleard et al 1984,b).

As Gilleard (1984) points out, "the measurement of strain is an unresolved area for research". Under these circumstances he feels that the results of different studies must be compared with caution.
At this point our conceptualisation of these phenomena is vague. In a colloquial sense these terms may be used synonymously to refer to the same thing. This review of various studies using different measures raises some doubt about this. Clearly this is an issue which needs to be addressed and clarified.

Whether these measures reflect the same underlying phenomena is uncertain. It may be that differences such as those found by Gilleard (1984) and Zarit (1982) reflect the fact that some measures are more sensitive than others. Alternatively they may reflect different facets of strain.

This seems more likely as the results of Greene et al's (1982) factor analysis suggest that strain is multi-dimensional (Gilleard 1984). At this point it is safest to interpret the various measures as reflecting on the carer's general level of psychological well-being.

In general the results of these studies indicate that the impact of care-giving does result in what I shall refer to (for simplicity) as strain. It also seems that the degree of strain varies from one carer to another. From a practical point of view it is important to identify the main factors which determine the degree of strain and to understand what mediating factors can ameliorate strain.
THE DETERMINANTS OF STRAIN IN CARERS

As care-giving involves many things, research has attempted to identify the particular factors which contribute most to the level of strain in carers. In this section I will be reviewing the literature about the determinants of strain. A useful framework for doing this comes from Clark & Rabowski (1983) who categorise the tasks of care-giving into four groups. These are:

- tasks concerning the direct care of the dementing person
- intrapersonal tasks, concerns, and the difficulties of the carer
- tasks concerning interpersonal ties with family members
- tasks concerning interaction with broader societal and health care networks.

I will use these categories as headings to structure this section although I have deviated somewhat from what Clark & Rabowski include in these.

The Demands Involved in Direct Care

Among the tasks of direct care listed by Clark & Rabowski (1983) are the performance of the 'activities of daily living' for the dementing person and coping with upsetting behaviour. These include a number of things such as helping the person to wash, eat, and dress as well as coping with incontinence, aggression, and wandering.

Various terms have been used to label this category of things. Gilleard (1984) considered the term 'stressors' but felt it was not appropriate as it implies that these things are stressing when in fact they may not be. He preferred the term 'daily hassles' but this too has the problem of implying that these things are a hassle when
all carers may not agree on this. They are also often referred to as 'problems' which they need not be. For example, Gilteard (1984) found that for most of the items on a 'problem check-list' some carers rated them as being 'no problem' for them.

This issue of terminology is important for both conceptual and practical reasons. We need to be able to distinguish between the event and the experience, i.e. we must differentiate between what actually happens to the carer and how they perceive and react to this. The practical importance of terminology concerns the outcome of studies of care-giving. As we shall see later, when you ask carers about 'problems' you find something different than if you ask them about 'behaviour disabilities and disturbance'.

I prefer to use the term 'demands' to refer to the events which face carers as opposed to the longer phrase 'disabilities and behaviour disturbances'. I also prefer this to the term 'task' which more accurately concerns the response related to the demand. I do not wish to assume or imply that events are demanding or stressful but simply that any event entails some demand of the person who reacts or responds to it. This is true whether the demand is perceived as positive or negative.

Research has looked at three aspects of the relationship between demands and strain in carers. These concern the nature of the demands, the overall level of demands, and the length of time the carer has been coping with them. The nature of the demands.
a) Common demands

The range of demands encountered by carers is extensive (cf. Sanford 1975, Greene et al 1982, Gilleard 1984) as all will not be mentioned here. Studies of demands have aimed to identify the most common demands as well as the most aversive.

The usual way of identifying the most common demands is to find out which are reported by the majority (i.e. 50% or more) of the carers. Some are consistently found to be commonly reported by carers. These include incontinence, sleep disturbance, proneness to falls, and demanding behaviour (e.g. asking constant questions).

However, there are considerable differences between studies. For example, unlike Sanford (1975), Gilleard et (1982) did not find that the majority of carers reported faecal incontinence or proneness to falls. Yet they did find that being unable to leave the dementing relative alone was common while Sanford did not.

Such variations might be due to differences in the samples concerning the severity of dementia. Generally speaking, the type and level of demands depends upon the severity of dementia (e.g. faecal incontinence is more likely with greater severity).

However, it is more likely that these variations are due to differences in the method of gathering the data. The way in which carers are asked about demands can affect the results. Gilleard (1984) found that asking about behaviour disability and disturbance produced different results from asking about problems. For example, only 29% reported 'temper outbursts' when asked about problems while 60% did so when asked about behaviour disability and disturbance.
By asked carers to distinguish between the occurrence of events and their status as problems, the list of demands which Gilleard (1984) found to be common was not radically different from those found by others. But this does emphasise that there is a difference between events and the way they are experienced by the carer. As Gilleard points out, "while a problem reflects a disability, a disability does not necessitate the existence of a problem".

b) The worst demands.

Research has also been aimed at identifying which demands are least tolerable or most aversive. Different methods have been used such as asking which "would need to be alleviated to restore a tolerable situation at home" (Sanford 1975) and asking carers to rate individual demands on a continuum from "no problem" to "a great problem" (Gilleard 1984; Gilleard et al 1984,a).

Studies generally find that the worst demands are sleep disturbance (nocturnal wandering), incontinence, demanding behaviour, and dependent behaviour (e.g. unsafe to leave unsupervised, need assistance to get in/out of bed or on/off the commode).

Again, there are contradictions. For example, whereas Sanford (1975) found that physical aggression was hard to tolerate, Gilleard (1984) did not find that many carers rated this as a "major" problem.

Regardless of this what is most interesting is that there is considerable variation between carers as to how aversive particular things are. No demand was felt to be
'great' by all carers. Nor was there a single demand which was felt by all carers to be 'no problem'. Hence, how aversive an event is does not depend upon its inherent quality, but the subjective quality which it has for that particular carer.

c) Demands which contribute to strain.

Most studies which seek to identify which demands are determinants of strain look at different types of demands rather than specific demands. This is done by factor analysing measures of demands (e.g. problem checklist) to identify clusters or categories of demands.

This does not necessarily make interpretation easier as was Gilleard et al's (1982) intention. The different categories of demands which have been identified are not obviously equivalent. For example, Gilleard et al (1982) found five compared to Greene et al's (1982) three while Gilleard (1985) found six. Those which seem to be common concern behaviour disturbance, dependency, demanding behaviour, and disability. Other categories are less universal such as incontinence, disturbance of gait, wandering, and impairment of self care.

The results of these studies suggest that the most stressful demands are dependency, disturbed behaviour, and behaviour which demands time and attention. Incontinence and impaired self hygiene are also particularly stressful.

There were some contradictory findings which make it difficult to summarise. For example, although Gilleard (1985) reported a significant correlation between behaviour disturbance and strain, neither Zarit et al (1980) or Gilleard et al (1982) found this to be the case. To complicate things, Gilleard et al's (1982) category of 'disturbance' which was not significantly correlated with strain contained
aggression, accusations, bad language, and temper outbursts, all of which occur in Greene et al's category of 'mood disturbance' which was correlated significantly with strain. Also, two of the four items in Gilleard et al's (1982) 'disability' category concerned incontinence and they did not find this category to be significantly related to strain whereas Gilleard (1985) did.

In an attempt to summarise, authors have tended to; rely on what they believe these demands imply for the carer (cf. Greene et al 1982; Gilleard 1984). Behaviour which seems to reflect apathy and a lack of interest in others is more likely to be stressful. Strain is also likely to result when they feel frustrated in their need to do something for their dependent relative who would otherwise sit doing nothing but deteriorating. So again it seems that it is the meaning which events have for the carer which is most important in determining whether they represent a source of strain.

2) The overall level of demands.

Gilleard et al (1984,a) found that the 'number of problems presented to (the carer) by their dependent relative' was the principal determinant of strain and burden in carers.

It is not clear how to reconcile this with the common finding that the severity of dementia is not correlated with strain (Zarit et al 1980; Zarit 1982; Greene et al 1982). One would expect there to be a greater number of problems with greater severity.

This may be because the number of problems and severity are assessed in different ways. Severity is usually determined by a formal clinical interview with the
dementing person while the number of problems is assessed by interviewing the
carer. As the latter is a more subjective measure it is likely to be affected by the
carer's level of strain i.e. more problems will be reported by those under greater
strain.

This possibility is supported by evidence in the literature. Gilhooley (1984) found
that carers viewed the dementing person as more impaired than did the day hospital
staff. Whether carers where over estimating or day staff were underestimating is
not certain. However, this does indicate that observer's assessments of the level of
severity is not unbiased. In his study of the impact of day care, Gilleard (1985)
found that carers who felt that attendance had benefited their dementing relative
showed a reduction in their level of strain and also reported fewer problems than
they had prior to attendance beginning. This suggests that reducing strain altered
the carer's perception of events.

3) The duration of care-giving.

One would expect that higher levels of strain would occur in those who had been
providing care for longer periods of time. However, the duration of care-giving has
not been found to be associated with the level of strain (Zarit et al 1980; Gilleard et
al 1984,a). In fact Gilhooley (1984) found better morale and mental health in carers
who had been care-giving for a greater length of time.

There are two possible explanations. Carers who are unable to maintain good
morale and mental health may be filtered out over time as they admit their relative
into hospital. Or it may be that carers become better at coping as they learn from
experience.
It does seem likely that some sort of adjustment process must occur as their relative will change and the level of demands will increase as deterioration continues with passing time.

**Intrapersonal Demands of Care-giving**

When we speak about intrapersonal demands we refer to the fact that the carer must cope with the effect which other care-giving demands have upon them. These may be physical as well as psychological.

The carer's own health can be a source of strain. They may experience weakness and fatigue (Sanford 1975; Barnes et al 1981) and some find it hard to tolerate such physical problems (Sanford 1975). Poor health was one of three factors (including the level of disturbing and demanding behaviour) which Gilleard et al (1984,b) found to be the primary determinants of high emotional distress in carers.

As mentioned earlier, care-giving can have psychological consequences which may themselves contribute to strain. Anxiety (e.g. about the dementing relative's safety) can be difficult to tolerate (Sanford 1975). Fears about further deterioration can be a source of considerable burden for carers (Zarit et al 1980). The relative's level of dependency can lead to feelings of isolation and disruption of social life which represents a major problem to some (Gilleard et al 1982).

Not uncommonly carers react to their circumstances with frustration, anger, or resentment. Often it is their relative's behaviour which evokes such feelings. The fact that many experience a sense of guilt (Barnes et al 1981) suggests that feelings such as these are a source of stress. If the carer is to cope with these they must alter their reaction to the events which provoke them. As we shall see later, it has
been suggested that this can be achieved by changing their perception of their relative.

**Interpersonal Tasks of Care-giving**

The sorts of tasks which Clark & Rabowski (1983) include in this category concern the relationship between the carer and other family members. I think it is also appropriate to include the relationship between the carer and their partner.

The importance of the relationship between the carer and other members of their family has been studied in terms of what is called 'social support'. The two parameters of this which have been analysed are the amount and the quality of social support.

The frequency of visits from family members is the usual means of measuring the amount of social support. It is generally found that strain is not correlated with the amount of social support (Zarit 1982; Gilhooley 1984; Gilleard 1984,a). However, Zarit et al (1980) did find a lower level of burden in carers who received more frequent visits from family members which led them to conclude that social support is the main influence determining carers' ability to continue to provide care.

The quality of social support has been determined by asking carers to rate their degree of satisfaction with the input from other family members (Zarit 1982) and by ratings of the interviewer (Gilhooley 1984).

These studies indicate that the quality of the social support is significantly correlated with carers' level of strain. Those who feel satisfied with input from family members show lower levels of burden and distress and higher morale.
Gilhooley (1984) concludes that it is the carer's reaction to the support they receive which is the most important determinant of their psychological well-being. This factor is apparently more important to women as Zarit (1982) found the quality of social support was a major predictor of burden for women but not for men.

Two aspects of the carer-partner relationship have been considered in relation to the level of strain in carers; the quality of the premorbid relationship and the quality of the current relationship.

Johnson & Williamson (1980) believe that the quality or the premorbid relationship is important. They argue that in the context of a poor previous relationship the current demands will create tension and evoke feelings of guilt about the past. In order to void or escape this stressful situation they feel that the carer will cease to provide care.

The contribution of this factor to strain in carers has been studied. A poor premorbid relationship has been found by Gilleard and his colleagues (1984,a; 1984,b; and 1985) to be associated with poorer mental health ('emotional distress') and greater burden and strain. However, Gilhooley (1984;) did not find that it was associated with either mental health or morale.

It seems likely that this difference in results is due to differences in the way the quality of the past relationship was measured. One major difference is that Gilleard used ratings made by the carer while Gilhooley's ratings were made by the interviewer. Another difference was that Gilhooley used only a single, overall rating while Gilleard used the ratings on six items (e.g. How often did you laugh and joke together?)
There is also a problem in the fact that both measures are retrospective. It may be that the current experience of coping with demands and a high level of strain colour the carer's recollection of the past.

Different aspects of the quality or the present relationship have been considered as important to strain in the carer.

Reciprocity is believed to be important by Johnson & Williamson (1980) who feel that the carer needs some sort of recognition and appreciation for what they are doing. Although they address themselves particularly to "child" carers (i.e. those caring for their dementing Parent); the importance of reciprocity for spousal carers has also been acknowledged.

Reciprocity is an important aspect of what Hirschfeld (1981) called 'mutuality'. She argued that the carer was likely to abandon care-giving if the relationship was not one of mutuality. I shall say more about this later.

Although the importance of reciprocity has been asserted it has received little formal attention in research. Horowitz & Shindelman (cited in Gilleard 1984, p.82) devised a measure of reciprocity based on credits earned from the past by the dependant but they did not find this to be associated with the carer's level of strain.

Another aspect of the carer-partner relationship which has been studied concerns the dementing person's current degree of role fulfilment as a husband or wife. Zarit (1982) found that the level of burden was greater when the dementing partner was no longer felt to be fulfilling this role, but this was only the case if the carer viewed the extent of role fulfilment as important.
The sex or the carer is another aspect of the carer-partner relationship that has been studied. It is generally found that strain is greatest amongst wives caring for a dementing husband (Gilhooley 1984; Gilleard et al 1984,a; Gilleard 1985) although it is not clear why this is the case. It may be due to gender differences in the behaviour or attitudes of carers or gender differences in the behaviour of dependants. Alternatively it may simply be because women are more likely to openly express the difficulties and strain they experience.

Demands Concerning Interaction with Social & Health Care Networks

In this category Clark & Rabowski (1983) include tasks involving interaction with helping professionals as well as acquiring knowledge and ability to use the system to their benefit. The aspects of this area which have been studied are the impact of day care services and the effect of input from various services on strain in carers.

As most of this concerns their ameliorating effects on strain they will be discussed in the next section. Although it is suggested that this factor can contribute to strain in carers there are no formal studies of this.

MEDIATORS OF STRAIN

What stands out from the literature reviewed so far is that strain is not simply a function of the level or nature of demands confronting the carer. What seems to be more important is how they react to the demands, i.e. how they perceive them. It is the meaning of events which is most important. Their level of strain depends upon how they construe what is currently happening to them and what implications this has for them about the past and/or the future.
This has important implications concerning the mediators of strain. Cognitive strategies will be very effective means of ameliorating the potentially stressful impact of events. Carers who perceive the experience of care-giving in certain ways are likely to experience less strain.

One example of this which has already been mentioned concerns the carer's perception of the support they receive from other family members. Those who feel this is satisfactory experience less strain than those who do not, regardless of the actual amount of support received.

The impact of day care provision is another mediating factor for which the carer's subjective assessment is more important than quantitative variables. The level of strain does not depend on how many times a week the carer's relative attends the day hospital (Gilhooley 1984). It depends more on the carer's perception of the effects of day hospital attendance.

Gilkeard and his colleagues (1984,c; 1985) have studied the relationship between the carer's expectations and perceptions of the day care service and the impact which it has on them. Essentially they found that the mediating effect of day care depends on the carer's willingness to maintain their relative's attendance which in turn depends upon the benefits they perceive this has for them and for their relative. Strain and emotional distress was ameliorated in those who expected that day care would be to their advantage but not to their relative's disadvantage.

The carer's understanding about what their partner is suffering from can also influence their level of strain. The importance of talking with them about the diagnosis and prognosis is emphasised by Rabins (1984). He feels that providing
information by explaining the diagnosis can help the carer understand their relative's behaviour and thus can ameliorate some of its stressful effects. Rabins also argues that explaining the prognosis and discussing the future with the carer can reduce their sense of uncertainty and rid them of unrealistic expectations. In so doing it may be possible to reduce the sort of fears which Zarit et al (1980) found to contribute to burden.

The carer's understanding of their relative's current behaviour is also important as it may have significant meaning for them. Some carers believe it is a form of retribution for the past (Sheldon 1982). Actions may be seen as deliberate attempts to inconvenience or upset the carer (Gilleard 1984) as often the carer feels their partner has acted intentionally (Barnes et al 1981). Such counterproductive attributions can contribute to strain but Rabins (1984) believes they may be avoided or modified by appropriate counselling.

It is also possible that carers will adopt adaptive cognitive strategies without receiving counselling. Various of these have been suggested in the literature.

Rather than an attribution which may heighten the emotional impact of events (e.g. "he is doing this on purpose") the carer may attribute them to the illness, thus redefining them as unintentional. It is also possible that they minimise the problems or even ignore them. Gilhooley (1980, cited in Gilleard et al 1982) found that this strategy was helpful in maintaining morale in some carers.

Denial of the illness and its ramifications is another common reaction, particularly in the earlier stages of deterioration (Barnes et al 1981). The beneficial effects of temporary denial in the early stages of severe life events has been argued elsewhere (cf. Lazarus & Golden 1981). Barnes et al (1981) feel that initial, temporary denial
can lessen the emotional impact of facing what is happening thus it may allow time for things to 'sink in' although if prolonged it may be inappropriate and counterproductive.

Johnson & Johnson (1983) have described how carers' perception of their dementing relative progressively changes as the dementia gets worse. One early stage is for the carer to change their expectations about their relative so they are in line with the lower level of ability and functioning.

As deterioration continued Johnson & Johnson noted what they called 'depersonalisation'. The carer's perception of their relative seemed to undergo redefinition until they were "no longer the person (I) know". They argued that such perceptual shifts were necessary to help the carer avoid negative emotions, such as guilt, and eventually enabled them to make the decision about hospitalisation easier.

Redefinition or the relationship is acknowledged as another means of coping with the burden of care-giving. Johnson & Catalona (1983) identify two major types of strategies for achieving this, "distancing" and "enmeshment". These strategies may be either essentially psychological or physical.

Distancing can be achieved psychologically in various ways. Ignoring things is one way which was mentioned earlier. Rationalisation is another method. For example, the carer may tell themselves, "I have done/am doing all I possibly can" and so distance themselves from the fact that their efforts have little impact. Gilleard et al (1982) believe that progressive disengagement from the emotional relationship while continuing to provide physical care is an effective coping strategy.
Role entrenchment is a psychological means of enmeshment. It involves the carer accepting the role of a full-time care-giver such that the exchange relationship is redefined. It then gains a new meaning in which the carer obtains a sense of reward for fulfilling the care-giving role. Hirschfeld (1981) points out that this role can be reinforced either internally (by the carer valuing their efforts) or externally (by the response of their relative or others to their efforts).

Gaining gratification from the care-giving relationship is an important part of what Hirschfeld (1981) called 'mutuality'. She identified three aspects of a relationship which define its mutuality. The carer must feel that care-giving is meaningful and feel gratified in such a relationship. They must also feel that the relationship is one of reciprocity. Hirschfeld felt that mutuality was crucial to continued care-giving as it would counterbalance the 'tension' which occurs. As long as mutuality was greater than strain the carer would continue to look after their relative.

It is worth noting that coping strategies are not only psychological. Physical distancing (Johnson & Catalona 1983) is one strategy which is more likely to be used by carers who do not live with the dementing person. It can be achieved by decreasing the amount of time spent with the relative and/or increasing the length of time in between visits. For carers living with their relative, it can be achieved by reducing the amount of direct contact or involvement in direct care activities (by distributing this to others such as a community nurse). Gilhooley (1984) has suggested that use of this strategy may account for better morale amongst male carers.

Physical enmeshment (Johnson & Catalona 1983) is a strategy linked with social regression. In this the carer reduces their involvement with others so they can devote the majority of their time to their relative. This strategy is particularly
common amongst spousal carers especially those who are childless or live far from their children. Although it may initially facilitate care-giving it is difficult to reverse and so becomes increasingly counterproductive as they become the sole, isolated care-giver.

It is clear from this review of the literature that psychological factors are important mediators affecting the impact which events will have on carers. It also seems that here are a wide variety of psychological strategies which may effectively help the carer to cope by adapting to the changes which take place.

However, at this point most of what we know about such psychological factors is based on speculation and clinical intuition. It is important to pursue these with related research. The question is, which factor or factors could most fruitfully be studied and how might this be done? In the next section I will consider both of these before going on to explain what was done and found in the present study.

**LINKING SECTION**

1) **Choice of psychological factors to study.**

A number of potential psychological factors that might be studied arise from what was said about mediating factors. These include denial, attributions, expectations, and minimisation to name but a few.

The main issue affecting the decision about which to focus upon is the question of how common any particular factor may be. Although denial may be used by some carers it seems unlikely that all or most carers will display this. The same applies to most of the other factors. For example, Gotay (1985) found that about 25% of
the women she studied who were confronted by a serious life event (cancer) did not attempt to make causal attributions about their experience.

This issue is important because it affects the scope of any potential research. In order to fruitfully study this topic it would be necessary to include a large number of carers in the sample and probably investigate numerous factors. This was not feasible in the case of the present project.

There are however some aspects of the care-giving situation which are common to all carers and which could be studied. These concern two facts; (1) the carer will have known their relative fairly intimately over an extended period, and (2) their relative will be doing things which are quite out of character from what they have usually been like.

a) Change of Image
Over the course of the relationship with their relative, the carer will have developed a conception or Image of them based on experience. They feel they know what sort of person their relative is and feel fairly confident that they know what to expect from them. When their relative does something which does not fit with their expectations or Image, they will either discount this as exceptional (e.g. to do with unusual circumstances) or modify their Image of them. Hence, over the years the carer's Image of their relative will have undergone gradual changes to maintain an acceptable fit with their experience of them. So long as this fit is maintained and their relative does not do anything which is grossly at odds with the Image, the carer will not experience a crisis in reconciling events with their expectations.
In the case of dementia though, it is likely that the carer's relative will do things which provoke such a crisis. For example they may become very forgetful and lose the ability to dress themselves or wire an electrical plug. Such changes may happen suddenly or only occur intermittently in the early stages.

Hence, the sequelae of dementia will be stressful for the carer as they are out of character and constitute a clash between their Image of their relative and their experience. This clash may be source of stress as events will have been unexpected. For this reason it can be predicted that the greater the discrepancy between their Image and what their relative is actually like at present, the greater will be the level of strain they experience. In cases where the carer has modified their Image in relation to the changes in their relative we would expect a lower level of strain. This would be due to events being less out of character so there would be less clash between their Image and their experience.

Comments made by others in the literature support these ideas. The effects of deterioration in dementia include personality change, reduction in task capacity, and inappropriate behaviour (Sheldon 1982). However, "behavioural disturbance and disability do not emerge as fully-fledged problems, but rather evolve from minor oddities and lapses" (Gilleard 1984). Such character changes are the most difficult symptoms for carers to reconcile (Barnes et al 1981). In order to cope they must "work through changes in the life long relationship" between themselves and their relative (Clark & Rabowski 1983). As time passes the carer adjusts in order to cope better (Gilhooley 1984).
b) Current role fulfilment

Another potentially fruitful area for research concerns the effect of the dementing process on the relative's role fulfilment. Related to the first fact mentioned above is the fact that the carer and their relative will have developed a certain role relationship. The sequelae of dementia will alter the extent to which the relative fulfils their role. In the case of the marital relationship, the reduction in role fulfilment can contribute to the burden of care-giving as Zarit (1982) found.

The spousal role has at least two main facets. One facet is task oriented and relates to who does what (e.g. cooking, paying bills, driving, house decorating, etc.). The other facet relates to personal characteristics (e.g. dominance, decisiveness, empathy, caring, etc.). Dementia affects both of these facets by reducing the capacity to fulfil tasks and by producing personality changes (Sheldon 1982).

The importance of fulfilling the task oriented facet of the spousal role has been studied by Zarit (1982). This was because her definition of "role" was restricted to this facet (role was defined as "who performs what functions in what relationship"). Hence the importance of role fulfilment in terms of personal characteristics is not presently known and will therefore be addressed here. We would expect that greater strain will be associated with the dementing partner having less of the personal characteristics which are felt to be ideal by the carer.

c) Reciprocation

As it has been suggested that reciprocation is an important aspect of the carer-relative relationship (Johnson & Williamson 1980; Hirscheld 1981; Barnes et al 1981; Gilleard 1984) it has been included amongst the personal characteristics
assessed in the measure of role fulfilment. We would expect that if the carer's partner is perceived as being less reciprocative than ideal, then their level of strain will be greater.

d) **Quality of the premorbid relationship**

The quality of the relationship prior to the onset of dementia has also been found to affect the strain in carers (Gilleard et al 1984, a, b). However, this has been studied only in a limited sense. Gilleard's measure included only six items which focused primarily on negative interactions (e.g. anger, arguments, possessiveness, and tension). Therefore, it would be interesting to study other aspects of the premorbid relationship. For example, the degree of previous role fulfilment in terms of the personal characteristics mentioned above. This has been done here with the prediction being that strain will be greater where the partner did not previously fulfil the ideal level of such personal qualities.

2) **Choice of measures.**

Having decided which factors to study the next question to address is what measures are to be used.

a) **Image change**

There are several possible techniques which might be used to measure the carer's Image of their partner. However, there are certain constraints. The measure should not be exceedingly long or difficult. The method of Q-sorts (cf. Cronbach 1970) was rejected for this reason as it involves the sorting of many statements and would require the carer to “give patient consideration to a long list of questions” (p.587).
Another important constraint is that the measure must be relevant to describing an elderly person in both normal and dementing states. Repertory grid technique (Fransella & Bannister 1977) meets both of these requirements. It can be relatively brief and its contents can be designed to suit the purpose.

Repertory grid technique was chosen to measure Image as well as past and present role fulfilment. One could be designed whose contents (bi-polar constructs) were relevant to all three factors and would include an item related to the issue of reciprocity.

The extent to which the carer had modified their Image of their partner was assessed by asking them to describe their partner in terms of the grid constructs both now and prior to the onset of dementia. The difference between these two grids would reflect the change in Image. The prediction was that strain would be greater amongst carers whose Image had changed less (i.e. lower difference scores). It was also predicted that constructs which related to the sequelae of dementia (e.g. memory disturbance, greater dependence) would tend to change whilst those not directly linked to the effects of deterioration (e.g. optimism, goodness) would tend to remain stable.

b) Level of behaviour disturbance

As I have suggested that Image will alter in relation to changes in their partner it is also important to have some measure of this variable. The most appropriate measure would be a behaviour rating scale designed for use with the dementing elderly. There are various 'problem' checklists and behaviour disturbance scales from which to choose (e.g. Gilleard 1984; Greene et al 1982).
I chose the behaviour and mood disturbance (B&M) scale used by Greene et al (1982) as it was readily available and at the time was the only scale which had been factor analysed to provide subscales of behaviour disturbance. According to the reasoning above, it can be predicted that if Image is modified in accordance with behavioural changes, then this scale should be correlated with the measure of Image change. It is also predicted that the B&M scale will be correlated with measures of strain and mental health as this was found to be the case by Gilleard et al (1984,a).

c) Severity of dementia

Because behaviour scales such as the B&M scale are based on the carer's perceptions of their partner it would be useful to know how accurate these perceptions are. As was mentioned earlier, strain has been found to correlate with the level of behaviour disturbance but not with the severity of dementia. It was also mentioned that the level of behaviour disturbance reported by the carer was less when their level of strain was lower. This suggests that the level of behaviour disturbance which is reported is influenced by the carer's level of strain.

For this reason a clinical measure of severity was included to investigate this relationship. The measure chosen was the Clifton Assessment Procedure for the Elderly (CAPE) (Pattie & Gilleard 1979). Only the Information and Orientation subscale (I/O) was used to minimise the demands placed on the dementing person. This subscale has been found to be a valid indicator of the severity of cognitive impairment (Pattie 1981).

We would expect to find as others have that severity, as measured by the CAPE, is not correlated with the measure of strain. We would also predict that CAPE scores
will not necessarily correlate with B&M scores if the latter depend more upon the level of strain than the degree of severity.

d) Type of dementia

Another variable which relates to the issue of Image change is the type of dementia. Although there are various different types the two most common are Alzheimer (DAT) and multi-infarct types. One of the diagnostic differences between them concerns the relative preservation of personality which is said to occur in multi-infarct dementia (Slater & Roth 1972).

If this is so we would expect less change in Image in carer's whose partner suffered with this type compared to DAT. We would also expect the nature of Image change to be different with change primarily concerning disability (e.g. memory problems, self reliance, activeness). To address this the scale devised by Hachinski et al (1975) was used to differentiate the two types of dementia.

e) Current role fulfilment

As mentioned above, repertory grid technique was also used to measure the extent of current role fulfilment. It is particularly useful for addressing the personal characteristic facet of role rather than task fulfilment.

As a measure of the partner's current characteristics would be obtained as part of the assessment of Image change it remained only to determine the extent to which this differed from the carer's notion of the ideal characteristics. This was measured by asking the carer to complete another grid using the same constructs but rating them in relation to their idea of what would be ideal in a husband/wife.
We would predict that the level of strain would be greater amongst carers for whom there was a larger discrepancy between the Ideal Husband/Wife grid and the grid describing their partner 'Now'. It can also be predicted that strain will be associated with a larger discrepancy on the construct concerning reciprocation.

f) Quality of the relationship in the past
Repertory grid technique was also used to measure the quality of the relationship prior to the onset of dementia. This was done by assessing the previous degree of role fulfilment in relation to the ideal characteristics in the same way current role fulfilment was measured. The difference was that previous fulfilment was determined by comparing the Ideal Husband/Wife grid with the grid describing what their partner had been like 'Before' the onset of dementia.

The prediction in this case was that greater strain would be associated with a larger discrepancy between these, two grids.

g) Strain and psychological well-being
As mentioned earlier, there are a variety of available methods for measuring the overall impact of care-giving. It was decided in this case to use the strain scale devised by Greene et al (1982) as it had been factor analysed to reveal different components of strain.

However, it was also felt important to place this measure in context by comparing it with other measures of psychological well-being.

The General Health Questionnaire (GHQ) (Goldberg 1978) was chosen as a measure of general mental health. It is a reliable and widely used measure that is not restricted for use with carers (as strain and burden scales are). The GHQ-28
was used as this form has subscales which might be useful for more detailed analysis of mental health.

A measure of self esteem was also used as it could be easily incorporated in the repertory grid measures. This was done by asking carers to complete a grid describing themselves as they are and as they would be ideally. Lower self esteem would be reflected in a larger difference between the two grids.

If the measure of strain is a valid measure of general psychological well-being we would expect it to correlate with both the GHQ and self esteem.

h) A comparison group

Most studies of carers do not include a comparison group. Hence, we can not always be sure to what extent, if any, carers differ from other elderly people who do not have a dementing partner.

For this reason, a comparison group was included in the present study. It was not possible to have them complete all the measures which the carers did as some were inappropriate. As a result they did not complete the B&M or strain scales, nor was their partner interviewed using the CAPE.

For the measures which they did complete we can make the following predictions. The degree of Image change will be less in the control group. It will also be of a different quality as the nature of the changes in their partner will be different from the changes in dementing partners.
We can also predict that the degree of current role fulfilment will be greater in the control group as their partners will have changed less than dementing partners. No difference in the degree of past role fulfilment would be expected.

If the consequences of care-giving do affect general psychological well-being in carers we would expect them to have higher scores on the GHQ and Esteem measures. Thus they would have poorer mental health and lower self esteem than the control group.

**DESIGN**

The present study included two groups of people (Dementia and Control) who were interviewed during the 11 month period from October, 1985 to February, 1986. Although it concerned married couples, only one of the partners was interviewed.

The interview included up to 8 different measures, however those in the Control group completed only six of these. Of these six, three concerned the person being interviewed and three required them to refer to their partner. All measures related to the present except one which was retrospective. In this the informant was asked to describe what their partner had been like in the past.

Two additional measures included in the study concerned the dementing partner of those in the Dementia group. These were completed by interviewing them and by asking their doctor to complete a questionnaire about their medical status.
It was possible to compare the two groups on the six measures they both completed. Relationships between the different measures were analysed within the two groups as well as within the combined groups (for the six shared measures).

**METHOD SECTION**

11) Informants.

A total of 110 couples were interviewed, with 20 in each group. The Dementia Group consisted of 12 men and 8 women who were the primary carer for their dementing partner. The Control Group consisted of 12 men and 8 women none of whose partner was suffering from dementia. In all cases only one of the spouses was interviewed to gain information about their partner. The age range, mean age, and duration of marriage for husbands and wives in each group is given in Table 1 below.

**TABLE 1**

<table>
<thead>
<tr>
<th>Demographic variables of informants in both groups.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dementia Group (Carers)</strong></td>
</tr>
<tr>
<td><strong>Men</strong></td>
</tr>
<tr>
<td>Age Range</td>
</tr>
<tr>
<td>Mean Age (overall)</td>
</tr>
<tr>
<td>Range of Marriage Duration</td>
</tr>
<tr>
<td>Mean Marriage Duration (overall)</td>
</tr>
</tbody>
</table>

Details of the informant's partner who was being described are given in Table 2.
TABLE 2

Demographic variables of informants' partners.

<table>
<thead>
<tr>
<th></th>
<th>Dementia Patient Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td><strong>Age Range</strong></td>
<td>61-86 yrs.</td>
<td>61-85 yrs.</td>
</tr>
<tr>
<td><strong>Mean age</strong></td>
<td>71 yrs.</td>
<td>76 yrs.</td>
</tr>
<tr>
<td>(whole group)</td>
<td>(73 yrs.)</td>
<td>(72 yrs.)</td>
</tr>
</tbody>
</table>

a) Carers

Carers were obtained from several sources in the Exeter and Torbay areas, primarily through staff of the community team for the elderly or the day hospital. The selection criterion were that they lived with their partner who had been diagnosed as suffering from a dementing process and that they should be the primary care-giver. An attempt was made to gather a group which represented a range of ages, levels of coping, and duration of care-giving.

As carers were not obliged to take part in the study they were to some extent self selecting. Only one whom I approached declined to take part. As most were initially approached by another professional who knew them it is possible that some declined prior to me contacting them, although these were unknown to me. Carers were excluded if staff felt the interview might adversely affect them. For this reason carers who were struggling with the early stage of accepting the diagnosis or who were under a particularly high level of stress tended to be excluded. Never the less, it was possible to interview some carers whose partner had only fairly recently been diagnosed as dementing and some under considerable stress.
The length of time over which carers had been providing care varied ranging from about 1 year to 10 years. As this factor has not been found to be a significant influence on strain it was not controlled for in the present study.

There were various differences amongst the carers concerning factors mentioned earlier which have been shown to influence strain in the carer. The level of social support ranged from having regular support from friends or family members living locally to having little or no support due to family and friends being far away or deceased. This is quite common in the Southwest where couples move into the area after retiring and become isolated from family and friends. In at least one case the family was a source of stress as they were critical and antagonistic towards the carer.

The level of professional help also varied amongst carers. In most cases (18 of the 20) the carer’s partner attended a day hospital or day centre, however this varied between one and three days a week. There were also differences in the amount of outside assistance and input carers received (e.g. home help, community nursing).

The patients who were being described by their spouse varied in terms of type, severity, and length of illness. An attempt was made to obtain a mix of DAT and multi-infarct type dementia. A range of the degree of severity was also sought although it was more difficult to find cases of mild dementia as these often have not come to the attention of the professionals, particularly where onset is gradual and insidious. The length of illness also varied considerably however in most cases it was only possible to approximate as it is very difficult to accurately pinpoint onset.
b) Control Informants

Control couples were obtained from several sources. The majority (N=111) were contacted through over-60's social clubs. Others were contacted through housing associations (N=11) and through carers (N=2). Very few who were approached declined to take part in the study (N=11). As all were initially approached by an intermediary who knew them (e.g. warden or club secretary) others may have declined unbeknown to myself.

The criterion for selection were age, health, and willingness to take part. An attempt was made to match the Control and Dementia groups in terms of sex and age of the person being described.

No couples were included in the Control group if the informant's partner was dementing. However, as one would expect of this age group, various physical and medical complaints had produced changes in some partners. These included diabetes, angina, arthritis, and impaired hearing.

2) Measures

A number of different measures were obtained using a variety of measurement tools. A list of measures and the tools used to obtain them is given in Table 3. Each will be described in a subsequent sub-section.
TABLE 3
Measures and measurement tools

<table>
<thead>
<tr>
<th>Measures</th>
<th>Measurement Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spouse's Image of their dementing partner (Now &amp; Before)</td>
<td>Repertory Grid</td>
</tr>
<tr>
<td>Degree of Image Change (Now vs. Before)</td>
<td>&quot;</td>
</tr>
<tr>
<td>Degree of partner's fulfilment of spousal-role (past &amp; present)</td>
<td>&quot;</td>
</tr>
<tr>
<td>Informant's Self Esteem</td>
<td>&quot;</td>
</tr>
<tr>
<td>Informant's Mental Health</td>
<td>General Health Questionnaire (GHQ-28; Goldberg 1978)</td>
</tr>
<tr>
<td>Carer's level of Strain</td>
<td>Strain Scale (Greene et al 1982)</td>
</tr>
<tr>
<td>Degree of Behaviour &amp; Mood Disturbance in the dementing partner</td>
<td>Behaviour &amp; Mood Disturbance Scale (Greene et al 1982)</td>
</tr>
<tr>
<td>Type of dementia</td>
<td>Ischemia Scale (Hachinski et al 1975)</td>
</tr>
<tr>
<td>Severity of dementia</td>
<td>Clifton Assessment Procedure for the Elderly CAPE (Pattie &amp; Gilieard 1979)</td>
</tr>
</tbody>
</table>

a) Repertory Grid

i) Design

A rating grid was designed which included 5 elements and 17 bipolar constructs. The elements were: Partner Now, Partner Before, Ideal Husband/Wife, Self, and Ideal Self.

All 17 constructs were supplied and were intended to fulfil the following criterion:

- they should be applicable to all 5 elements (i.e. they should fulfil Kelley's (1959) range of convenience corollary)
- some would be particularly relevant to the personal qualities of the informant's partner
- of these, some would be intended to address the prediction that
• certain qualities are more likely to "change" as a result of the dementing process while certain other qualities are less likely to change (i.e. they will "hold")

• some of the constructs would be particularly relevant to aspects of the spousal role fulfilled by the informant's partner

• some of the constructs would be particularly relevant to the personal qualities of the informant.

The constructs included in the grid used in the study were derived from several sources. These included psychological literature about the elderly and dementia (e.g. Gilleard 1984), assessment of personality change following organic brain damage (Moffit 1978), and semantic differential measures of personal meaning (Osgood 1957). Several of the constructs were based on constructs elicited from elderly informants during a preliminary pilot study. This will be described now.

ii) Pilot study

Prior to doing the main study, a pilot study was done with the primary aim of designing a useful Grid measure.

A total of 14 people were seen during the pilot study. Most (13) were recruited by approaching residents in various local housing associations. All of them were a member of an elderly couple where neither was suffering from dementia. It was more difficult to locate couples where one member was dementing. The one who was found one was living with her demented husband in a residential home where they had recently moved.

Of the 111 who were interviewed during this phase, 5 were men and 9 were women. Their ages ranged from 60-89 years with the mean age being 711 years
old. The ages of the partner being described by these people ranged from 69-89 years with the mean being 76 years old.

The procedure used in completing the various measures with people during the pilot study were the same as those used during the main study (see Procedure section) except for some minor differences which will be described now.

As well as supplying constructs, several were from each person using different methods described elsewhere in the repertory grid literature (cf. Fransella & Bannister 1977). The method of triads was initially used but abandoned as the elderly informants found it difficult. The method of dyads, which involves comparing two people, was more successful (e.g. “In what way are Jim and Tom like/different from each other?”).

A total of 36 bi-polar constructs were elicited in this way. Several which were used by more than person were selected out for use as the basis of some of the constructs in the final form of the grid. In addition to the original form of the Grid, two modified versions were designed. Copies of Grids I, II, and III can be found in Appendices 1, 2, and 3.

Grid II contained 11 changes from Grid I. Two of these involved constructs for which people were reluctant to one of the poles. It seems reasonable to assume that for reasons of social desirability people were hesitant to refer to their partner as either "unhelpful" or "dishonest". Hence, these constructs were changed to "helpful-helpless" and "open-secret". Two new constructs were also added as they quite often occurred in some form in the constructs which had been elicited. These were "proud-modest" and “patient-impatient".
TABLE 4
Constructs changed in Grids II and III
to reduce the problem of social desirability

<table>
<thead>
<tr>
<th>Grid II</th>
<th>Grid III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tough-Tender</td>
<td>Tough hearted-Tender hearted</td>
</tr>
<tr>
<td>Dominant-Submissive</td>
<td>Gets their own way-Goes along with others</td>
</tr>
<tr>
<td>Passive-Active</td>
<td>Inactive-Active</td>
</tr>
<tr>
<td>Good-Bad</td>
<td>Good-All right</td>
</tr>
<tr>
<td>Selfish-Generous</td>
<td>Thinks of own needs-Thinks of others needs</td>
</tr>
<tr>
<td>Calm-Excitable</td>
<td>Doesn't get ruffled-Gets ruffled</td>
</tr>
<tr>
<td>Helpful-Helpless</td>
<td>Helpful-Not helpful</td>
</tr>
<tr>
<td>Modest-Proud</td>
<td>Modest-Not modest</td>
</tr>
</tbody>
</table>

Grid III contains 11 changes from Grid II. Most of these concern the issue of social desirability mentioned earlier. The constructs which were changed for this reason are shown in the Table 4.

The modified constructs were generated in various ways with the intention of retaining the same basic meaning. The alternative wordings were derived from the dictionary and thesaurus as well as ad hoc. These were “tested-out” on colleagues to assess the extent to which they had the same meaning as the initial form and to which they were not socially undesirable.

Two additional constructs included in Grid III were derived from constructs which had been elicited. These concerned "optimism" and "contentedness".
It is worth noting that many of the elicited constructs were similar to some of the original supplied constructs. This suggests that the supplied constructs were relevant and suitable for this study.

The final construct which was added related to the issue of reciprocity in the marital relationship. Initially the wording used was "Reciprocates- Doesn't Reciprocate" but it was immediately clear that these words were too esoteric. Hence the wording used in the final form of Grid III was "Gives as much in return as they get from others-Doesn't give as much as they get".

A list of the 17 constructs used in the final form of Grid III can be found on the sample grid recording form in Appendix 3.

iii) Grid measures

A measure of the person's current Image of their partner was obtained by asking them to rate the element "Partner Now" on each of the 17 constructs.

A retrospective measure of their previous Image was obtained from their ratings for the element "Partner Before". This related to their partner prior to the onset of dementia (for the Dementia group) or to a period approximately 10 years previously (for the Control group).

A measure of the degree of Image Change was obtained by comparing ratings for the element "Now" with the element "Before".

Current role fulfilment was obtained by comparing ratings on the element "Now" with the element "Ideal Husband/Wife". Previous role fulfilment was found by comparing ratings on "Before" with ratings on "Ideal Husband/Wife".
The informant's level of self esteem was measured by comparing ratings for the elements "Self" and "Ideal Self".

b) General Health Questionnaire

The General Health Questionnaire (GHQ) used was that devised by Goldberg (1978). Of the different forms of the GHQ which are available, the 'Scaled GHQ' (GHQ-28) was chosen as it is more suited for research where psychiatric 'case-ness' is not as relevant as obtaining a more general measure of mental health. An example of the GHQ-28 can be found in Appendix 4.

The GHQ-28 consists of 11 subscales which are based on factor analytic studies of questionnaire items. The subscales are; Somatic Symptoms / Anxiety & Insomnia / Social Dysfunction / Severe Depression.

The usual method of scoring (0-0-1-1) was used. With this scores on the subscales can vary from 0-7 with the total varying from 0-28.

Goldberg (1978) has found the reliability of the GHQ to be good (test-retest reliability of .90; split-half r=.95).

c) Strain Scale

The measure of strain used was that devised by Greene et al (1982). This consists of 15 questions which are answered by the carer to indicate the frequency with which they experience each item. For example;

"How often do you feel that there will be no end to the problem?"
All the time/ Frequently/ Sometimes/ Rarely/ Never.

The form used in this study was the same as that used by Greene (personal communication).

Greene et al's factor analysis of this questionnaire revealed three clusters of items. Hence the subscales are;

   Personal Distress / Domestic Upset / Negative Feelings.

Each item is scored from 0 to 11 hence subtotals can be found for each subscale and a total Strain score can be calculated for the whole scale.

Greene et al report a test-retest reliability for the entire scale of .85 with reliability coefficients for individual subscales ranging from .72 to .88. An example of the Strain Scale can be found in Appendix 5.

d) Behaviour & Mood Disturbance Scale

This measure was also devised by Greene et al (1982) and contains 31 items. Each item refers to the dementing partner and is answered by the carer to indicate the frequency of occurrence. For example;

   "How often does ____ wander outside the house at night?"

   All the time / Frequently / Sometimes / Rarely / Never

The form used in the present study was the same as that used by Greene (personal communication).

Greene et al's factor analysis revealed three subscales. These are;
Apathetic-Withdrawn / Active-Disturbed / Mood Disturbance

Scoring is similar to the Strain scale with the range for each item of 0-11. It is also possible to obtain scores for each subscale as well as a total B & M score.

Test-retest reliability for the entire B & M Scale was found by Greene et al to be .811 with reliability coefficients for individual subscales ranging from .73-.90. A copy of the B & M Scale can be found in Appendix 6.

e) Ischemia Scale
The Ischemia Scale was devised by Hachinski et al (1975). Its purpose is to identify patients who are likely to be suffering a multi-infarct type dementia. The scale contains 17 items concerning the patient's medical state and history. A list of these can be found in Appendix 7.

Items receive a score of either 2 or 1 depending upon their relative weight in contributing to the diagnosis. The total score can range from 0-18. Patients whose total score is 7 or greater are considered to be likely cases of multi-infarct dementia.

f) Clifton Assessment Procedure of the Elderly (CAPE)
The CAPE (Clifton Assessment Procedure of the Elderly) (Pattie & Gillear 1979) can be used to measure cognitive and/or behavioural disability in the elderly. In the present study only the Information/Orientation (I/O) subscale was used.

The I/O subscale (see Appendix & ) contains 12 questions which assess the person's orientation in time and place. Each item counts as 1 point with the total
score ranging from 0-12. The lower the total score, the greater the severity of cognitive disability.

The test-retest reliability of the I/O subscale has been found to be good \( (r=.811) \) by Pattie & Gilleard (1979).

**PROCEDURE**

1) **The Dementia Group**
   Carers were interviewed in their own homes. It took one to one and a half hours to complete all measures (excluding the CAPE) in the following order:

- the Hachinski Scale was partially completed by gathering general medical background (e.g. concerning onset of dementia, past medical complaints, etc.)

- the B & M Scale was completed concerning their partner's present state

- the Strain Scale was completed in the context of understanding the effects this state of affairs had on the carer

- the Grid was completed beginning with ratings for Now and followed by ratings for Before, Self, Ideal Self, and Ideal Husband/Wife

- the GHQ was completed

The B & N and Strain Scales were both done in the same manner. For each item, carers were asked whether they experienced or observed that particular thing. If so they were asked how often it happened. The frequency categories which apply to
that particular item were stated as part of the question to allow them to indicate
the frequency of occurrence. For example;

"Does he ever hoard useless things?.....How often does he do that? all the time?
frequently? sometimes? or rarely?"

When it was apparent from previous comments that the item in question had
occurred the carer was only asked how often it happened. For example;

"How often do you get cross and anger with your husband ?. ... All the time?
frequently? sometimes? or rarely?"

In presenting the Grid to carers, they were not told explicitly that it was intended
to indicate their past and present Image of their partner. Instead it was couched in
terms of indicating what they were like generally as a person at present and prior
to the onset of dementia.

Ratings for the element Now were obtained by providing each construct word-pair
in the context of the question of what their partner was presently like (e.g. "Is your
wife a tough hearted or tender hearted sort of person ")). Carers usually indicated
one of these although they were told that in some cases they might feel that their
partner was neither definitely one or the other, but in between the two. They were
also offered the possibility that a particular construct might be inapplicable, which
was sometimes the case.

If in their reply the carer did not state the extent to which their partner was like
that pole of the construct (e.g. "Oh, he's fairly tender hearted.") then they were
asked to indicate whether he/she was "very", "fairly", or "just a bit" like that. The
three alternatives were presented in random order across items. They were not necessarily all three presented if the carer's non-verbal behaviour (e.g. tone of voice or expression) suggested the degree. In that case only two of the alternatives were mentioned (i.e. "fairly or very", "fairly or just a bit"). Their choice between these was usually quite definite but if not they were reminded of the third alternative.

A rating of 1 for example if they indicated "very tough hearted" Carer's responses for each construct were recorded as indicated below:

1 for "very tough hearted"
2 for "fairly tough hearted"
3 for "a bit tough hearted"
4 for "neither one or the other but about half way"
5 for "a bit tender hearted"
6 for "fairly tender hearted"
7 for "very tender hearted"

The ratings from 1-7 correspond to the spaces on the record sheets (see Appendix 3) from left to right. If the carer indicated that a particular construct was not applicable this was noted and dealt with accordingly when analysing the data. (N.B. Constructs were only ever felt to be inapplicable to the element Now, not on any of the other elements.)

When carers were uncertain about how to reply (e.g. "Well that depends on the situation, sometimes he's one and sometimes he's the other.") they were asked to answer in general terms (e.g. "Well what is he like in general"). This usually prompted a definite reply. When there was a question about what the construct
poles meant the carer was offered a general definition. Such a definition was prepared for each construct so that the same definition was offered consistently across all carers who enquired (e.g. optimism/pessimism was defined as the tendency to either look on the bright side or take a gloomy view of things). It was rarely necessary to provide a definition as carers usually responded when asked to generalise.

To obtain ratings for the element Before, carers were asked to describe what their partner had been like prior to the onset of the dementing process. They were asked to identify a point in time prior to onset when their partner had been "as good as they ever were". To help them focus on that time they were asked questions to describe where they were living then, with whom, and what each of them had been doing (e.g. work). They were frequently reminded that their description referred to this time by the phrasing of questions (e.g. “At that time was he ....”, “Ten years ago was he..........”).

To obtain ratings for the element Self, carers were asked to describe themselves in terms of the 17 constructs. The purpose of this was presented as being to place the description of their partner in context by indicating what their perspective was. All carers did this willingly without hesitation. As with earlier ratings they were first asked which of the two poles best described them and then to indicate the degree.

To complete ratings for the element Ideal Self the 17 constructs were gone through again but this time the carer was asked to indicate what they would ideally like to be like with respect to each construct. They were constantly reminded that they were describing their ideal rather than actual self. This was done by prefacing each question with a reminder such as "Would you prefer to be ______ or?", or
"Would you rather be _____ or _____?" A similar format was used when asking about the degree e.g. "Would you ideally be fairly, very, or just a bit like that?"

When the carer was uncertain about their reply they were asked to generalise. If, for example they felt that their answer depended upon situational factors they were asked to indicate what they were like or would like to be like in general. If they were still unsure they were reminded of the option that they were "neither strongly one or the other" or they were "half way between the two". If this was the case a rating of 11, the middle point of the rating scale was recorded. In practice this occurred infrequently as most carers gave a definite response when asked to generalise.

The final element which carers were asked to complete was that of Ideal Husband/Wire. This was intentionally set apart from the elements Now and Before to help them differentiate their ideal from their actual spouse. Additional steps were taken to make this distinction. Carers were asked to describe what their notion of an ideal husband or wife (the opposite of their role) was and not what their partner was or had been like. It was suggested that this might be like their actual partner in some ways but might be different in other ways. They were told that everyone had their own individual idea about what was ideal and part of the present study was interested in finding out about this.

To help establish a cognitive set different from that concerning their actual Spouse, carers were presented with an imaginary scenario and then asked to rate each of the constructs as regards the ideal husband/wife. The scenario was as follows:

"Imagine that an old friend of yours who has been single all their life comes to you and announces that they intend to get married. They say that they are going to begin looking around but they are not quite sure what it is
important to look for in selecting a partner. Knowing that you have had years of experience of married life they have come to you for advice about what an ideal husband/wife would be like. What would you say to them? Would the ideal husband/wife be ___ or ____?“.

With this lead-in each construct was taken in turn. Carers were constantly reminded that their rating concerned an ideal husband/wife. This was done by prefacing questions with a reminder such as "Would the ideal husband be ___ or _____?" or "Would this ideal sort of husband be ___ or ____?" The same was done concerning the degree e.g. “Would they ideally be very, fairly, or just a bit like that?”

As when rating earlier elements, if the carer was uncertain about their response they were asked to generalise. When still unsure they were reminded that their reply might be "neither strongly one or the other" or "half way between the two", in which case the relevant rating was recorded. If the carer referred to what their partner was like they were reminded that the element in question was the ideal not their actual spouse. If their reply to any question was that the ideal was just what their partner was like they were asked what that was and this was confirmed as being their ideal e.g. "So the ideal husband would be ____?"

The final part of the interview with the carer involved them in completing the GHQ. Unlike the preceding measures, the items on the GHQ were read by the carer rather than being asked by the interviewer. There were one or two exceptions to this when the carer asked that the items be read to them because of difficulty with reading. In these cases each item was read aloud and the four alternative replies were stated (see Appendix 4).
Carers were asked to give an indication of how their health had been recently by completing the GHQ. With the form in front of them the instructions and the first item with its alternative responses were read to them. They were left to complete the remaining items on their own. If they had a question they were reminded of the instructions and the item and associated responses was read to them for their reply.

The I/O subscale of the CAPE was completed by interviewing the dementing person either at the day hospital or in their own home (if they were not attending day hospital). Eight were seen by myself and nine were to be seen by one of the staff at the day hospital. It was not possible to complete the remaining three as one died, another had a serious stroke, and in the case of the third the carer was unwilling to allow his wife to be interviewed. It was intended to complete the CAPE as near in time as possible to the interview with the carer but this was not always the case.

2) The Control Group
The procedure used in interviewing those in the Control group was similar to that used with carers except that only the Grid and GHQ were completed. They were told that the researcher was interested in finding out in what ways people changed as they get older. They were told this was being done by asking someone who knew them well (i.e. their spouse) to describe what they are like at present and what they were like at some point in the past. The elements of the Grid were presented in the same manner as described above except for the element Before. The GHQ was also administered as described above.

In a few cases both spouses were interviewed however the data from only one of them was used in the analyses. The decision about which partner to interview or
which partner's data to use in the analysis (if both were interviewed) was made according to what was necessary to match members in the Control group with members of the Dementia group in terms of the sex and age of the person being described. In one or two cases where both partners were interviewed the decision about whose data to use was determined by the fact that some complicating factor (e.g. deafness) tainted the validity of that informant's responses.

As the person being described by the Control group subjects was not dementing, a different method of presenting and completing the element \textit{Before} was used. The person was asked to choose and focus upon a time in the past about 10 years earlier. They were asked to identify this point in time by fixing it to a particular event (e.g. a holiday or moving house). If the event which they chose seemed to have been a period of relatively unusual or significant impact (e.g. the disruption of moving house; the transition to retirement) they were asked to shift their focus backwards or forwards in time to a more usual, stable period. As was done with the carers, they were asked questions to elaborate upon the details of that time as a means of facilitating the task of describing their partner as they were then. They were also constantly reminded of this task by phrasing questions as described earlier.
RESULTS SECTION

The data was analysed using SPSS-X (Statistical Package for the Social Sciences) packages for t-test and Spearman's correlation. Analyses were done between groups (i.e. comparison of Dementia and Control groups) as well as within groups. The emphasis will be placed on results concerning the Dementia group and will be compared with those concerning the Control group where appropriate.

Tables and correlation matrices containing the results of all the main analyses have been included in the appendices for reference. Comparison of the two groups on various measures was done using t-tests and the results of these are given in Appendix 9. The relationships between measures for the combined groups are shown in the correlation matrix in Appendix 10. Correlations between measures within the Control group can also be found in the matrix in Appendix 10 while the correlation matrix for the Dementia group is in Appendix 11.

Correlations were also done within the Dementia group between various measures (Strain, GHQ, Esteem, and B&M) and separate constructs. Correlations concerning IC constructs are shown in Appendix 13. Those concerning RFB and RFN are in Appendices 111 and 15.

Throughout this section asterisk symbols will be used to indicate significance levels. These will be:

** p<0.01  * p<0.05

Analysis of Descriptive Data

The Dementia and Control groups were compared with each other in terms of the ages of the informant (e.g. carer) and the person being described (e.g. dementing partner) as well as the length of marriage. As Table 5 shows, there was no
significant difference between groups on any of these variables, therefore the two
groups were matched in terms of these general demographic variables.

TABLE 5
Comparison of Groups on Demographic Variables

<table>
<thead>
<tr>
<th></th>
<th>Dementia Group</th>
<th>Control Group</th>
<th>T</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Age of Informant</td>
<td>69 yrs.</td>
<td>72 yrs.</td>
<td>-1.02</td>
<td>0.31</td>
</tr>
<tr>
<td>Mean Age of Partner</td>
<td>72 yrs.</td>
<td>72 yrs.</td>
<td>0.30</td>
<td>0.77</td>
</tr>
<tr>
<td>Mean Marriage Duration</td>
<td>43 yrs.</td>
<td>41 yrs.</td>
<td>0.48</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Analysis of Change of Image

Change of Image (IC) was analysed between groups as well as within groups. The
comparison between groups concerned differences in total IC as well as differences
in terms of the rank ordering of constructs from largest to smallest mean change
scores. The correlations between IC and the other measures was done for both
Dementia and Control groups as well as for the combined groups.

1) Comparison between groups of total IC

There was a significant difference in total IC between groups. The mean for the
Dementia group was 119.65 (s.d. 18.65) compared to 13.115 (s.d. 8.12) in the
Control group. The two groups were also compared in terms of which constructs
changed the most and which changed the least. This was done by calculating the
mean change score for each of the 17 constructs (using the absolute difference
between the ratings for Now and Before) and then rank ordering these within each
separate group (See Appendix 12). The difference between rank orderings was
analysed with Spearman’s Rank Correlation Test.
The correlation between the two sets of ranks was not significant (rho= .19, p>.10). Hence, the constructs which tended to show the most change in the Dementia group were different from those showing the most change in the Control group. The four which tended to change the most in the Dementia group were Memory, Self Reliance, Activeness, and Reciprocation. In the Control group these were Dominance, Calmness, Memory, and Reservedness.

2) Relationship between IC and other measures
Within the Dementia group IC was significantly correlated only with RFN and the level of behaviour & mood disturbance (B&M). The correlations between IC and other measures in this group were;

<table>
<thead>
<tr>
<th></th>
<th>B&amp;M</th>
<th>RFN</th>
<th>RFB</th>
<th>GHQ</th>
<th>Strain</th>
<th>Esteem</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC</td>
<td>.46*</td>
<td>.82***</td>
<td>-.26</td>
<td>.26</td>
<td>.26</td>
<td>-.02</td>
</tr>
</tbody>
</table>

RFN was also significantly correlated with IC in the Control group (r=.75*** ) and in the combined groups (r=.93***).

The correlation between IC and GHQ was significant overall (r=.51*** ) although it did not reach the level of significance in either of the separate groups (r=.23 in the Control group).

RFB was significantly correlated with IC in the Control group (r=.116* ) which was not the case in the Dementia or combined groups (r=.18).

3) Test of predictions about 'change' & 'hold' constructs.
A Mann Whitney test was used to determine whether constructs which were predicted to 'change' would tend to be high in the rank order of mean change scores while constructs predicted to 'hold' would be lowest in the rank order.
The position of these constructs in the rank order is shown in Appendix 12. This shows that in broad terms the prediction was upheld. However, this failed to reach the level of significance (U=9, p>.05).

11) **The relationship between separate constructs of IC and measures psychological well-being**

Correlations were done between the separate constructs of IC and total scores for strain, GHQ, and Esteem (see Appendix 13). In order to analyse the relationship between the measures and the size of the difference (e.g. between their rating of Caring for Now and Before), absolute differences were used in the correlations. Signed differences (i.e. -6 to +6) were used to analyse the relationship between measures and the direction of the difference (e.g. whether strain relates to an increase or a decrease in Caring).

a) **The relationship between strain and constructs of IC**

The level of strain was significantly correlated with the size of the difference in three of the IC constructs; Goodness (r=.55**), Contentedness (r=.119*), and Caring (r=.39*).

These indicate a relationship between higher strain and greater changes in Goodness, Contentedness, and Caring. There was not a significant correlation between strain and the degree of change in Reciprocation (r=.21).

Strain was related to the direction of change in the case of only one construct, Modesty (r=.60*). In this case greater strain is associated with a reduction in Modesty.
b) The relationship between GHQ and separate IC constructs

There was a significant correlation between the size of the difference in Goodness ($r=.53**$) and total GHQ. There were also significant correlations between GHQ and the direction of change in two IC constructs; Modesty ($r=.61**$) and Caring ($r=-.53*$). Poorer mental health is associated with reduction of modesty and caring.

c) The relationship between Esteem and separate IC constructs

Esteem was significantly correlated with the degree of change in Goodness ($r=.113**$). There were no significant correlations between Esteem and the direction of change in any of the constructs.

In summary, there is a difference in the degree and quality of change in Image (IC) between the Control groups. Image changes to a greater extent among the carers in the Dementia group. The nature of the change is also different in this group, particularly concerning reciprocation. This construct is amongst those which tend to change the most in the Dementia group while it is amongst those which change the least in the Control group.

Contrary to prediction, strain was not significantly correlated with the total degree of Image change (IC).

However, there were significant correlations between some separate constructs of Image and strain. Strain is higher amongst those who perceive their partner as being less modest than they were before. Greater strain is also associated with bigger changes in caring, contentedness, and goodness. Despite the fact that Reciprocation tends to change a great deal it was not significantly related to carer's level of strain.
There is a significant overall correlation between GHQ and IC although this did not reach the level of significance in the Dementia group alone. Overall then, poorer mental health was associated with a greater degree of change in Image. Amongst carers in particular, mental health is worse if the partner is perceived as being less modest and less caring than before. Poorer mental health is also associated with a larger degree of change in goodness.

The level of self esteem amongst carers was not found to be significantly correlated with the total degree of IC. However, lower self esteem was associated with a larger degree of change in goodness.

The degree of IC is significantly correlated to the total B&M score such that a larger change in Image is associated with a higher level of behaviour and mood disturbance in the dementing partner. IC is also significantly correlated with RFN (in the combined and separate groups) such that the greater the change in Image, the lower the degree of current role fulfilment.

Analysis of Role Fulfilment Before (RFB)

1) Comparison between groups of RFB
There was no significant difference in total RFB scores between the two groups (t=1.73, p=.09). The mean RFB score in the Dementia group was 15.60 (s.d.= 8.1) compared to 20.115 (s.d. = 9.5) in the Control group.

The qualitative difference in RFB between the two groups was analysed as it was for IC. The discrepancy between the Ideal and previous role fulfilment was found by calculating the difference between the ratings for Before and Ideal Partner on
each construct. Mean discrepancy scores were then rank ordered within each group (see Appendix 16). A significant correlation was found between the two sets of rank orders using Spearman's rank correlation test ($r=.67**$).

This indicates that there is no significant qualitative difference in RFB between the Dementia and Control groups. The constructs which tend to show the largest discrepancy are similar in both groups. The same is true of the constructs which tend to show the smallest discrepancy.

2) Relationship between RFB and other measures

Within the Dementia group RFB was significantly correlated only with Esteem and GHQ. The correlations between RFB and other measures in this group were;

<table>
<thead>
<tr>
<th></th>
<th>Esteem</th>
<th>GHQ</th>
<th>IC</th>
<th>RFN</th>
<th>Strain</th>
<th>B&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>RFB</td>
<td>.57***</td>
<td>.42*</td>
<td>-.26</td>
<td>.24</td>
<td>.19</td>
<td>.10</td>
</tr>
</tbody>
</table>

Esteem and GHQ were also significantly correlated with RFB in the combined groups ($r=.39**$ and $r=.37**$ respectively) but not in the Control group.

Although the correlation between RFB and RFN did not reach significance in the Dementia group it did in the combined groups ($r=.41**$) and in the Control group ($r=.60**$). IC was significantly correlated with RFB only in the Control group ($r=.46*$).
3) **Relationship between separate RFB constructs and measures of well-being**

a) **Separate RFB constructs and Strain**
Total strain was significantly correlated with the size of the discrepancy between Ideal and past role fulfilment in the case of four RFB constructs; Contentedness ($r=0.57^{**}$), Helpfulness ($r=0.42^{*}$), Goodness ($r=0.40^{*}$), and Caring ($r=0.40^{*}$).

The direction of the discrepancy in RFB constructs and strain was significantly correlated only for Reservedness ($r=0.47^{*}$). In this case higher strain was associated with the partner being less reserved in the past (i.e. more Outgoing) than the Ideal.

b) **Separate RFB constructs and GHQ**
Total GHQ was significantly correlated to the degree of discrepancy (between role fulfilment in the past and the Ideal) in four cases; Contentedness ($r=0.79^{***}$), Memory ($r=0.53^{**}$), Patience ($r=0.40^{*}$), and Calmness ($r=0.46^{*}$).

There were significant correlations in the direction of discrepancy of three RFB constructs and GHQ; Contentedness ($r=0.65^{**}$), Memory ($r=0.46^{*}$), and Calmness ($r=0.50^{*}$). Higher GHQ I scores were associated with a poorer than Ideal memory and being less calm and less contented than Ideal.

c) **Separate RFB constructs and Esteem**
There were significant, positive correlations between the degree of discrepancy in five RFB constructs and Esteem; Memory ($r=0.54^{**}$), Patience ($r=0.58^{**}$), Optimism ($r=0.55^{**}$), Contentedness ($r=0.63^{**}$), and Calmness

Esteem was significantly correlated with the direction of the discrepancy for three of these constructs; Memory ($r=0.49^{*}$), Patience ($r=0.45^{*}$), and Optimism ($r=0.48^{*}$).
Higher Esteem scores were associated with being less patient, less optimistic and more forgetful than Ideal.

In summary, there was no significant difference in the total degree of role fulfilment before (RFB) between the Dementia and Control groups. In addition, there was no significant difference between groups in the quality of...In general, the constructs for which there was the largest discrepancy between the Ideal and actual partner were similar in both groups. The same can be said of the constructs which showed the least such discrepancy.

Although the correlation between strain and the total degree of past role fulfilment (RFB) is not significant there are significant correlations with some of the separate constructs. Greater strain is associated with larger discrepancies in the past from the Ideal levels of contentedness, caring, helpfulness, and goodness. It is also associated with the partner having previously been less reserved than ideal.

The correlation between GHQ and RFB was significant in the combined groups but not in the Control group alone. These two measures were significantly correlated in the Dementia group such that poorer mental health was associated with a greater discrepancy between the degree of role fulfilment in the past and the Ideal. Amongst carers, poorer mental health was associated with the partner having been less contented, less calm, and more forgetful in the past than their Ideal partner would have been. The greater these discrepancies, the poorer their mental health. Poorer mental health was also associated with a larger discrepancy in patience.

Esteem is also significantly correlated overall with but within the Control group this correlation did not reach the level of significance. It did however in the
Dementia group such that lower self esteem is associated with there having been a larger discrepancy in the past between the actual role fulfilment of the partner and the carer's ideal. In particular, self esteem was lower if the partner had previously been less patient, less optimistic, and more forgetful than the Ideal. The greater the discrepancy in these qualities, the lower the level of self esteem. The degree of discrepancy concerning calmness and contentedness was also positively correlated with the level of self esteem in carers.

**Analysis of Current Role Fulfilment (RFN)**

1) **Comparison between groups of RFN**

There was a significant difference between groups in the degree of current role fulfilment ($t=8.17^{***}$). The mean RFN in the Dementia group was 54.45 (s.d. = 17) compared to 19.25 (s.d. = 9.1) in the Control group.

The qualitative difference in RFN between the two groups was assessed in the same way as it was for RFB (see above).

The rank order of constructs in RFN according to mean discrepancy score (between their partner at present and their Ideal) can be found in Appendix 17.

The similarity between the two sets of ranks was analysed using Spearman's rank order correlation test. The groups were not found to be significantly correlated ($\rho=.06$). On the whole the constructs for which the discrepancy was greatest were different in the two groups. In the Dementia group these were Memory, Self
Reliance, Selfishness, Confidence, and Reciprocation. In the Control group they were Dominance, Patience, Optimism, and Memory.

2) Relationship between RFN and other measures

In the Dementia group RFN was significantly correlated with IC, B&M, GHQ, and Strain. The correlations between RFN and other measures in this group were:

\[
\begin{array}{ccccccc}
& IC & B&M & GHQ & Strain & Esteem & RFB \\
RFN & - .82*** & .62** & .53** & .39* & .34 & .24 \\
\end{array}
\]

The correlation between IC and RFN was also significant in the Control group (r=.75*** and the combined groups (r=.93***).

The overall correlation between RFN and GHQ was significant (r=.63*** but in the Control group this correlation did not reach significance (r=.31, p<.10).

Although the correlation between RFN and RFB was not significant in the Dementia group it was in the Control group (r=.60*** and in the combined groups (r=.41**).

3) The relationship between separate RFN constructs and measures of well-being

a) Separate constructs of RFN and total strain

Total strain was significantly correlated with the degree of discrepancy (between Ideal and present partner) for two of the constructs in RFN; Contentedness (r= -.55**) and Reciprocation (r=.53**). The correlations between strain and Memory (r=.37, p=.056) and Goodness (r=.37, p=.054) just missed reaching the level of significance.
Strain was significantly correlated with the direction of the discrepancy for four RFN constructs; Secrecy ($r = -.63^{**}$), Caring ($r = .57^*$), Contentedness ($r = -.45^*$), and Reciprocation ($r = -.46^*$). In these cases, higher strain was associated with greater Secrecy, less Caring, less Reciprocation, and less Contentedness than the Ideal partner.

b) Separate RFN constructs and total GHQ

Total GHQ was significantly correlated to the degree of discrepancy in three RFN constructs; Contentedness ($r = .64^{***}$), Reciprocation ($r = .42^*$), and Goodness ($r = .40^*$).

The direction of the discrepancy was significantly correlated with GHQ in four constructs; Caring ($r = .70^*$), Modesty ($r = -.59^*$), Contentedness ($r = -.58^{**}$), and Secrecy ($r = -.56^*$). In these cases GHQ was correlated with the partner being less Caring, less Modest, less Contented, and more Secret than the Ideal.

c) Separate RFN constructs and Esteem

There was a positive, significant correlation between Esteem and the degree of discrepancy in Goodness ($r = .42^*$) and Memory ($r = .44^*$). Esteem was also correlated with the direction of the discrepancy in Memory ($r = .53^*$) and Caring ($r = .51^*$). In these cases Esteem was associated with the partner being less Caring and having a poorer Memory than Ideal.

In summary, there are quantitative and qualitative differences in RFN between the two groups. The degree of current role fulfilment is significantly less (i.e. the discrepancy is larger) in the Dementia group. The nature of discrepancy between Ideal and current partner is also different. In particular the discrepancy is larger
amongst the carers concerning self reliance, selfishness, reciprocation, and confidence.

The correlation between Strain and RFN was significant amongst carers such that the less their partner fulfils the ideal role standards, the higher is their level of strain.

The level of strain is also significantly correlated to particular aspects of RFN. In the Dementia group greater strain is associated with the partner being less contented and less reciprocative than the ideal partner. The larger the discrepancy in these qualities, the greater the strain. Greater strain is also associated with the partner being less caring and less open (i.e. more secret) than ideal.

The overall correlation between RFN and GHQ was significant (although this did not reach significance in the Control group). Hence, poorer mental health was associated with a larger discrepancy in the extent to which the partner currently fulfilled their role.

In the Dementia group particular aspects of current role fulfilment related to the carer's mental health. It was worse if their partner was less contented, less open, less caring, and less modest than ideal. Poorer mental health was also associated with a bigger discrepancy from the ideal levels of contentedness, reciprocation, and goodness.

Amongst carers, total RFN was not significantly correlated with Esteem (nor was it in the Control group). However, lower self esteem was associated with the partner being more forgetful and less caring than ideal. It was also related to bigger discrepancies in forgetfulness and goodness. RFN was significantly correlated with
B&M such that the greater the level of behaviour and mood disturbance, the less the degree of current role fulfilment.

IC was also significantly correlated with RFN overall and in both groups. The greater the change in Image, the less the extent of current role fulfilment.

Analysis of Strain

The level of strain was only measured amongst carers in the Dementia group. Correlations were done between total strain scores and the other measures. In addition, the relationship between the three separate subscales of the strain questionnaire and the other measures was analysed using Pearson's correlation test.

1) The relationship between total strain and other measures
Total strain was significantly correlated with Esteem, and B&M.

<table>
<thead>
<tr>
<th></th>
<th>GHQ</th>
<th>Esteem</th>
<th>B&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain</td>
<td>.65***</td>
<td>.39*</td>
<td>.38*</td>
</tr>
</tbody>
</table>

An analysis of multiple regression was done with all other measures and Strain as the dependent variable. This revealed that GHQ was the best predictor of Strain (beta=.65**).

2) The relationship between strain subscales and other measures
The correlations between the separate subscales of the strain measure and the other measures are shown in Table 6. Personal Distress was significantly correlated with GHQ and Esteem. Domestic Distress correlated significantly with GHQ. Negative
Feelings significantly correlated with GHQ, Esteem, B&M, RFB, and RFN. None of the subscales correlated significantly with IC.

**TABLE 6**

Correlations between Strain Subscales and Other Measures in the Dementia Group

<table>
<thead>
<tr>
<th>Strain Subscales</th>
<th>Personal Distress</th>
<th>Domestic Upset</th>
<th>Negative Feelings</th>
</tr>
</thead>
<tbody>
<tr>
<td>GHQ</td>
<td>.59**</td>
<td>.44*</td>
<td>.61**</td>
</tr>
<tr>
<td>Esteem</td>
<td>.42*</td>
<td>.12</td>
<td>.46*</td>
</tr>
<tr>
<td>B&amp;M</td>
<td>.15</td>
<td>.33</td>
<td>.55**</td>
</tr>
<tr>
<td>RFN</td>
<td>.20</td>
<td>.28</td>
<td>.59**</td>
</tr>
<tr>
<td>RFB</td>
<td>.14</td>
<td>-.07</td>
<td>.55**</td>
</tr>
<tr>
<td>IC</td>
<td>.11</td>
<td>.32</td>
<td>.23</td>
</tr>
</tbody>
</table>

In summary, strain was significantly correlated with GHQ and Esteem such that higher levels of strain were associated with lower self esteem and poorer mental health.

All three subscales of strain were significantly correlated with the general level of mental health. However, self esteem was significantly correlated with only two of the strain subscales (i.e. Personal Distress and Negative Feelings).

There is a significant correlation between strain and B&M such that higher levels of behaviour and mood disturbance are associated with greater strain. Correlations between strain subscales and B&M reached the level of significance only in the case of Negative Feelings.

**Analysis of the GHQ**

The GHQ was completed by all 40 subjects hence there are analyses for both Dementia and Control groups as well as for the combined groups.
There was a significant difference in mean total GHQ between groups. In the Dementia group the mean was 7.55 (s.d. = 6.18) compared to 1.95 (s.d. 2.4) in the Control group.

1) The relationship between GHQ and other measures
In the Dementia group total GHQ is significantly correlated with Strain and Esteem.

<table>
<thead>
<tr>
<th>GHQ</th>
<th>Esteem</th>
<th>Strain</th>
<th>B&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.80**</td>
<td>.65**</td>
<td>.32</td>
</tr>
</tbody>
</table>

The overall correlation between GHQ and Esteem was significant (r=.51***). However, this did not reach the level of significance within the Control group (r=.27).

A multiple regression analysis was done with all measures using GHQ as the dependent variable. This revealed that Strain and Esteem together were the best predictors of GHQ (beta values of .40* and .64** respectively).

2) The relationship between subscales of the GHQ and other measures
The correlations between GHQ subscales and other measures within the Dementia group are shown in Table 7. All 4 subscales were significantly correlated with total Strain, Esteem, and RFN. The correlations between RFB and all GHQ subscales except one (Social Dysfunction) were significant. Total B&M was significantly correlated with the Social Dysfunction subscale.
Correlations between GHQ subscales and other measures in the Dementia group.

<table>
<thead>
<tr>
<th>GHQ Subscales</th>
<th>Somatic Symptoms</th>
<th>Anxiety &amp; Insomnia</th>
<th>Social Dysfunction</th>
<th>Severe Depression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain</td>
<td>.64***</td>
<td>.66***</td>
<td>.53**</td>
<td>.48</td>
</tr>
<tr>
<td>Esteem</td>
<td>.83***</td>
<td>.66***</td>
<td>.65***</td>
<td>.71***</td>
</tr>
<tr>
<td>RFN</td>
<td>.47*</td>
<td>.40*</td>
<td>.51*</td>
<td>.52*</td>
</tr>
<tr>
<td>RFB</td>
<td>.38*</td>
<td>.46*</td>
<td>.23</td>
<td>.42*</td>
</tr>
<tr>
<td>B&amp;M</td>
<td>.18</td>
<td>.26</td>
<td>.46*</td>
<td>.27</td>
</tr>
<tr>
<td>IC</td>
<td>.20</td>
<td>.16</td>
<td>.33</td>
<td>.25</td>
</tr>
</tbody>
</table>

In summary, there is significant difference in mean GHQ score between the two groups with carers having poorer mental health. Although there is a significant correlation between GHQ and Esteem overall this only reached the level of significance in the Dementia group. GHQ was also significantly correlated to strain in this group. In carers then, there is an association between poor mental health, low self esteem and higher levels of strain.

Analysis of Esteem

Esteem was measured in both groups. There was no significant difference between groups (t = -.21). In the Dementia group the mean Esteem score was 18.4 (s.d. = 9.2) and in the Control group it was 19.0 (s.d. = 9.1).
1) Relationship between Esteem and other measures

In the Dementia group Esteem is significantly correlated with GHQ and Strain.

<table>
<thead>
<tr>
<th>Esteem</th>
<th>GHQ</th>
<th>Strain</th>
<th>B&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-.80**</td>
<td>.39*</td>
<td>.15</td>
</tr>
</tbody>
</table>

The correlation between Esteem and GHQ is also significant in the combined groups (r=-.51***) however, it did not reach significance in the Control group (r=.27).

In summary, there is no significant difference in mean Esteem between the two groups. Carers in the Dementia-group do not have significantly lower self esteem than those in the Control group. There is a significant correlation between Esteem and GHQ overall, however this correlation reached significance only in the Dementia group. Amongst carers lower self esteem is associated with poorer mental health.

Analysis of Behaviour & Mood Disturbance (B&M)

The level of behaviour and mood disturbance was measured only in the Dementia group. It was subjected to the same sorts of analyses described in the section about strain.

The level of behaviour and mood disturbance was significantly correlated only with total Strain.

<table>
<thead>
<tr>
<th>B&amp;M</th>
<th>Strain</th>
<th>GHQ</th>
<th>Esteem</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-.38*</td>
<td>.32</td>
<td>.15</td>
</tr>
</tbody>
</table>
The correlations between separate subscales of the B&M checklist and other measures is shown in Table 8. The only correlations to reach levels of significance are those between the Apathetic-Withdrawn and Active-Disturbed subscales and IC and RFN.

In summary, B&M and Strain are significantly correlated such that higher levels of behaviour and mood disturbance are associated with greater strain in carers. As none of the correlations between B&M subscales reached significance it does not seem that a specific aspect of behaviour and mood disturbance produces this relationship.

**TABLE 8**

Correlations between B&M subscales and other measures

<table>
<thead>
<tr>
<th>B&amp;M Scales</th>
<th>Apathetic-Withdrawn</th>
<th>Active-Disturbed</th>
<th>Mood Disturbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>RFN</td>
<td>.70***</td>
<td>.49*</td>
<td>.20</td>
</tr>
<tr>
<td>IC</td>
<td>.50*</td>
<td>.42*</td>
<td>.07</td>
</tr>
<tr>
<td>Strain</td>
<td>.35</td>
<td>.29</td>
<td>.30</td>
</tr>
<tr>
<td>Esteem</td>
<td>.26</td>
<td>.05</td>
<td>.02</td>
</tr>
<tr>
<td>GHQ</td>
<td>.28</td>
<td>.26</td>
<td>.25</td>
</tr>
<tr>
<td>RFB</td>
<td>.14</td>
<td>.01</td>
<td>.02</td>
</tr>
</tbody>
</table>

**Other Analyses.**

It was not possible to analyse the contribution of the severity of dementia as measured by the CAPE. Scores on the CAPE Information & Orientation scale were only available for 12 of those in the Dementia group. It was also not possible to study the effect of the type dementia as determined with the Hachinski Scale.
Only 4 of those in the Dementia group were identified as cases of Multi-infarct type dementia. The majority were cases of the more common form (i.e. Alzheimer’s type).

The possibility of sex differences was investigated using t-tests on all measures in the combined and separate groups. No significant differences were found. Hence, wives who cared for dementing husbands were not found to report higher levels of behaviour & mood disturbance or to be under more strain, have poorer mental health, or lower self esteem than husbands who were carers.

**DISCUSSION**

1) **Image change**

As predicted, the degree of change in Image was greater amongst carers compared to those in the control group. The nature of carers’ Image change was also different. In general the characteristics directly related to the sequelae of dementia (i.e. “change” constructs) tended to change the most as predicted. Therefore, it seems that most carers do take such changes in their partner into account and modify their Image accordingly. This is supported by the fact that greater Image change is associated with higher levels of behaviour disturbance.

The prediction about which constructs would tend to remain stable (i.e. “hold”) was not as clearly upheld. In particular optimism' and 'secrecy' changed more than expected. This may be due to the fact that these were constructs which a number of carers felt were no longer applicable hence they received the maximum change score.
Contrary to prediction, the extent of Image change in carers was not significantly correlated with their level of strain. In fact the correlation was opposite to the predicted direction. This was also true for the relationship between Image change and mental health. Overall (i.e. in the combined groups), people whose Image of their partner has changed most tend to have poorer mental health.

It seems that people are aware of changes which occur in their partner (whether this is related to dementia, retirement, or other factors) and that the extent of these changes is related to their mental health. This may be because such changes also mean that their partner has shifted further away from what they would ideally like them to be like. This is supported by the fact that in both groups greater change in Image is associated with a lower level of current role fulfilment.

2) Current role fulfilment

As predicted, the degree of current role fulfilment was correlated with strain and mental health in the Dementia group. Strain was greater and mental health was worse amongst the carers who perceived the largest reduction in their partner’s role fulfilment.

This result compliments the findings of Zarit (1982) and indicates that it is not only the task related aspects of role which are important, but also the personal characteristics of the marital partner.

The significance of reciprocity as suggested in the literature (Johnson & Williamson 1980; Hirschfeld 1981) is also born out by the present study. Strain was greater in carers whose partner was less reciprocative than they would like. Mental health was worse where the discrepancy from the ideal level of reciprocity was largest.
The importance of reciprocity is supported by the correlations between the measures of psychological well-being and the constructs related to caring and openness. When the partner is less caring and less open than ideal, strain is greater and mental health is worse. The carer's self esteem is also lower when their partner is less caring than ideal. It is likely that this reflects the importance that positive regard from a loved one has for one's own sense of self value.

Role fulfilment is not important only to carers. There was a significant correlation overall between mental health and current role fulfilment. Although this correlation did not reach the level of significance in the Control group it was in the same direction as in the Dementia group. In seems that in general, a person's mental health is worse when their partner is less like the sort of partner they would like to have. This makes sense and is consistent with clinical knowledge as well as the literature on marital satisfaction (e.g. Winch 1971).

The significant correlation between current and past role fulfilment in the combined groups suggests that the discrepancy between the ideal and actual partner widens with passing time. Thus it seems that partners tend to become less, I rather than more like the Ideal as they change. This means that the couples for whom role fulfilment was lowest in the past are also the ones for whom the level is lowest at present.

However, there appears to be a difference between groups concerning this phenomena. The correlation between RFN and RFB was not significant in the Dementia group whereas it was in the Control group. Therefore it seems that unlike those in the Control group, the carers for whom role fulfilment was lowest in the past are not necessarily the ones for whom it is lowest at present.
This is most likely due to the fact that the degree and nature of the changes in dementing partners is quite abnormal. This overshadows the usual trend and disrupts the rank order of role fulfilment which tends to be maintained in couples where neither spouse is radically affected by illness or other events.

The fact that self esteem was not significantly correlated with current role fulfilment in either group is at odds with what might be expected. As self esteem depends to some extent upon the role which a person fulfils and this is influenced by the role which their partner fulfils, we would expect one person's role fulfilment to affect the other's self esteem.

This may be because self esteem depends more upon the task oriented facet of role than the personal characteristics facet. Fulfilling role related tasks would enhance a sense of personal efficacy, competence, and control and thereby help to maintain a favourable self-regard. In a task-complimentary relationship such as marriage, the tasks fulfilled by one's partner could limit or overload their own role and thereby affect self regard.

However, the extent to which one fulfils their own ideal personal characteristics is not so dependent upon the degree to which their partner fulfils the characteristics of the ideal husband or wife. Hence the latter would not have such a strong affect upon self esteem. The exception to this would be personal characteristics such as caring which imply that the other is worthy of value. When their partner is less caring than they would wish they may begin to doubt their own value, hence such qualities would be related to self esteem as was the case amongst carers.
3) Previous role fulfilment

There was no significant difference between groups in the degree of previous role fulfilment. This suggests that the quality of the marital relationship in the past was no better or worse for carers than it was for those in the Control group.

Contrary to prediction, the degree of role fulfilment in the past was not found to be significantly correlated with the level of strain in carers. Such a finding is inconsistent with Gilleard et al's (1984,a) finding that greater strain and burden were associated with a poorer premorbid relationship.

This may be because the percentage of female carers was larger in Gilleard's sample (73% compared to 60% in the present study). Zarit's (1982) study suggests that the quality of the premorbid relationship tends to be a more important factor for women than men on certain measures of psychological well-being. She found that this factor was one of the best predictors of burden in wives, but not husbands. But she did not find a difference between sexes in the importance of this factor in relation to a measure of mental health. Hence it is possible that a larger proportion of female carers in the sample will distort the relationship between certain measures of psychological well being and, the quality of the premorbid relationship.

Significant correlations were found in the present study between the degree of past role fulfilment and the measures of self esteem and mental health. Self esteem was lower and mental health was worse the lower the degree of role fulfilment was in the past.

This is generally consistent with Gilleard et al (1984,b) who used the same measure of mental health but a different measure of the quality of the premorbid
relationship. However, it is not consistent with Gilhooley (1984) who used different measures of mental health and premorbid relationship.

The variations in results between these different studies (including the present one) suggests that the choice of measures will affect the outcome. In particular, the way in which the quality of the premorbid relationship is assessed varies widely. Zarit (1982) simply asked carers to rate their previous relationship on a 10 point scale from poor to excellent. Gilheard et al (1984,a,b) used ratings on 6 questions which primarily concerned interactions (e.g. How often did you....laugh and joke together? ....have upsetting arguments?). Gilhooley (1984) used a single rating by the interviewer (on a 5 point from very poor to very good) based on answers to unspecified questions.

A major weakness with these measures and the one used here is that they are all retrospective. We can not be certain to what extent the quality of the current experience affects recollection of the past. Ideally one would have actually measured the quality of the premorbid relationship prior to the onset of dementia, but this is impractical. However, a retrospective measure would be strengthened by corroborative information from other friends or family members. Indeed, this is a method which has been used to assess the quality of marital relationships (cf. Blood 1962).

It seems unlikely that the measure used here gives a completely inaccurate representation of the quality of the premorbid relationship. If the degree of previous role fulfilment depended upon the quality of the current situation we would expect a high, positive correlation between RFN and RFB. But this was not found in the Dementia group, hence to some extent the measure of RFB is unaffected by current circumstances.
A second weakness in these measures is the fact that they are limited to certain aspects of the marital relationship. This has many facets which include sexual, social, parental, economic, and domestic tasks. Multiple criteria of quality are more valid than a single criteria (Bell 1971, p.296) hence most measures will only address a particular aspect of the quality of the relationship. This issue should be remembered when interpreting all of these measures of the premorbid relationship. Satisfaction with the partner's role fulfilment is only one factor (Bell 1971) thus the degree to which the partner fulfilled the ideal gives only a partial reflection of the quality of the past relationship.

Certain aspects of RFB were found to be particularly important. Carers' psychological well-being related to the extent to which their partner had met the ideal in terms of being caring, helpful, and able to remember things. These seem to concern the relationship function of fulfilling needs for nurturance which is an important aspect of marriage (Blood 1962, p.556). This supports Johnson & Williamson's' (1980) suggestion that greater tension will occur when the premorbid relationship was poor. If in the past the carer had felt that their needs were not met they may feel ambivalent about meeting their dementing partner's needs. The dilemma of feeling torn between duty and resentment would adversely affect their general psychological well-being.

Patience and calmness were also important aspects of RFB which probably have some bearing on the number of arguments or disagreements in the past. These would have been more numerous with a more volatile (less calm) partner which would affect the balance of positive and negative experience. The fact that psychological well-being was worse in carers whose partner had been less calm
than ideal suggests that the demands of care-giving were more stressful in the context of less positive past relationship.

Thus, RFB is a valid measure of the quality of the premorbid relationship in so far as it reflects how positive and how nurturing the marital relationship was in the past.

4) The measures of psychological well-being

In general, all three of the measures which were used seem to be valid indicators of psychological well-being. Amongst carers, the measure of strain was significantly correlated with the measures of mental health and self esteem. The correlation between the measure of mental health and self esteem was also significant in the Dementia group and in the combined groups. However, there was some variation in the relationships between these and the measures of IC, RFN, and RFB.

Image change (IC) was the only measure on which there was consistency amongst the measures of psychological well-being. Neither strain, self esteem, or mental health was significantly correlated with IC.

However, changes in particular characteristics may affect the carer's psychological well-being. Understandably this is true if the partner becomes less caring than they were previously. This is also the case when the partner becomes less modest than they had been. As this concerns socially appropriate behaviour (not only physical modesty) it is likely that this relates to social strain (e.g. embarrassment).
The degree of current role fulfilment (RFN) was significantly correlated with strain and mental health but not self esteem. An explanation for this has already been given (i.e. self esteem is more dependent upon the task oriented aspect of role fulfilment). Hence, it seems that dementia reduces the extent to which the partner fulfils the personal role characteristics which are felt to be ideal which in turn adversely affects the spousal carer's psychological well-being.

We may understand this within the context of Exchange Theory (cf. Eiser 1980). This states essentially that the outcome of a relationship depends upon the balance of costs and rewards. The relationship will only be satisfactory if rewards outweigh costs. Dementia within the marriage constitutes an increase in costs (i.e. the demands of care-giving) with a reduction in the rewards associated with the partner's role fulfilment. Thus exchange theory implies that the relationship will become less satisfactory and may eventually become unsatisfactory.

According to this theory, whether or not a person remains in an unsatisfactory relationship depends upon whether there is an alternative which would be more satisfactory. Therefore, someone will remain in a relationship (e.g. decide not to place their partner in an institution) with costs outweighing rewards if the alternatives are not felt to be any better. Understandably, such circumstances are not conducive to psychological well-being.

The degree of previous role fulfilment (RFB) was significantly correlated with self esteem and mental health but not strain. We may be able to understand this by looking at the relationship between RFB and the separate subscales of the strain scale.
The degree of previous role fulfilment was significantly correlated with the 'negative feelings' subscale. The items in this subscale concern feelings of embarrassment, frustration, and anger. This makes sense considering what was said earlier about the importance of the nurturing and positive qualities of the relationship in the past. In the context of these having been less than ideal in the past, the carer is more like to react negatively to their partner's demands. It seems unlikely that this aspect of the past relationship would affect the items in the "domestic upset" subscale (e.g. changes in the household routine and standard of living). Yet one would expect this to affect the 'personal distress' subscale as its items concern the carer's tolerance of the current situation (i.e. feeling unable to cope, in need of a break, and depressed by the situation). In the context of a poor premorbid relationship one would expect the carer to be less tolerant. It is difficult to understand why this was not the case.

One general conclusion that can be drawn from these results is that the various measures of psychological well-being are not exactly comparable, each relates to a particular aspect of well-being. This is important to the study of different factors related to care-giving as these will most likely affect only certain aspects of well-being. Therefore, it is inadvisable to rely on a single measure. In addition, we must beware of generalising the results from one measure to another.

A final point relating to this issue concerns the importance of using measures which allow comparison with other groups. Authors have tended to assume that carers must be under greater strain and in poorer mental health than others their age, yet most (including myself) are amazed at how well they cope with their circumstances. Because of this it seems important to find out how much and in what ways carers' well-being differs from others. In this light it is interesting to note that the present study found that carers' mental health does indeed tend to be
worse than that in other elderly people however, the difference varies from none to considerable.

5) Constructs of particular importance.
Looking across the measures of Image and role fulfilment it appears that certain constructs are particularly important in relation to psychological well-being. Those which stand out by being correlated with at least two of the well-being measures are Goodness, Contentedness, Caring, Reciprocation, Memory, and Openness.

Caring, reciprocation, openness, and the ability to remember things can sensibly be grouped together as qualities which are important to the rewarding value of a relationship. If these are lacking then the relationship can offer little of positive value to the spouse. How this would influence psychological well-being has been discussed above in relation to exchange theory. This leaves us to consider the significance of Goodness and Contentedness.

It seems most likely that Goodness represents the "evaluative" dimension of a person's "semantic space" which Osgood et al (1957) identified. Their factor analyses of many semantic differentials (which are highly similar to the type of rating grid used here) revealed that this dimension was very general and central to peoples' perceptions of things. Hence, most constructs relate to this dimension and will therefore contribute to their evaluative judgement about an object or person. For example, whether they perceive a person as calm or reserved will have implications for them about how good or bad that person is.

Thus it is probable that a carer's rating of their partner in terms of Goodness is a distillation of their ratings on numerous other constructs. Essentially it reflects the implications which the other construct ratings have for their evaluative judgement
about their partner. This follows from the "organisation" corollary of personal construct theory (cf. Fransella & Bannister 1977) which states that constructs are organised in an hierarchical structure with subordinate and superordinate constructs being linked by implications.

Carers' psychological well-being was found to relate to changes in their partner's Goodness as well as to the discrepancy between the degree of Goodness they would like in a husband or wife and the extent to which their partner did not fulfil this (both in the present and past). There are three possible reasons why this has happened.

One straightforward explanation is that carers feel bad about construing their partner in a negative light. This may arise from a sense of commitment and loyalty (e.g. "They don't deserve to be seen by me in such terms."). This would have negative implications for themselves (e.g. "If I am capable of this I am not a good person"). The guilt which is commonly felt by carers may reflect this.

A related possibility is that carers' sense of self value (Goodness) may depend to some extent on the perceived value of their partner. Identification with another person or thing is one way in which people determine their own value. Hence it may be that when the carer perceives their partner's value depreciating or as being less than ideal their own self value is reduced. This is supported by the fact that self esteem was found to be lower in these cases. The evaluative dimension has been shown to be important to peoples' psychological well-being by Endler (1961) who found that improvement following psychotherapy was associated with changes in the evaluative component of the patient's self concept.
A third possibility relates to systemic theories of marriage (cf. Bodin 1981). These believe that families (or marital dyads) constitute systems which are homeostatic.

Hence, in order to be maintained as a viable system, compensatory changes will occur to counterbalance other changes which threaten the system. Because of this a member of the system may adopt a position which benefits the system even though it is against their best interests as an individual. Hence it is possible that the care-giving partner will "sacrifice" themselves by suffering a reduction in psychological well-being. This would be more likely to occur in symmetrical relationships (cf. Sager 1981).

One explanation of the importance of Contentedness relies on the psychodynamic notion of "projection" (cf. Sandler et al 1973). This refers to the projection of one's own feeling state onto another person. Usually this concerns unacceptable feelings, hence the externalisation of these is believed to act as a psychological defence against the adverse consequences of acknowledging that feeling in oneself.

Acknowledging a personal sense of discontent (with the partner or the marriage) may have adverse consequences such as a sense of guilt. Therefore, the carer may project this onto their partner to avoid or minimise such consequences. Greater discontent would be associated with poorer psychological well-being and perception of the partner as discontented.

A more parsimonious explanation would be that the partner actually is discontented and that this affects the carer's psychological well-being. A considerable sense of responsibility for the other's well-being is likely to exist ~ the marital relationship as emotional gratification is one o~ its important functions (Winch 1971). Therefore, when one's partner is unhappy or discontented this will
have implications for their spouse such as being a failure in terms of this marital function. Thus, to some extent the carer's psychological well-being will depend upon their partner's.

6) Clinical implications

The present results support clinical impressions that the mental health of carers of the elderly dementing is poor and is worse than it is amongst those without such responsibilities. However, some cope better than others thus the degree of psychological well-being varies considerably amongst carers. Assessment of this is important and the methods used here were found to be useful for this purpose.

As the quality of the premorbid relationship was found to be related to carers' psychological well-being some assessment of this will help to identify those most at risk. When the previous relationship was less nurturative and less positive than the carer would have liked it is likely that the current demands will adversely affect them.

Assessment of the current relationship may also be useful as it can contribute to strain and poor mental health. An important aspect of this is the degree to which this relationship is rewarding for the carer. This will be aided by gaining some indication of how caring, open, and reciprocative they perceive their partner.

Repertory grid technique may effectively be used in assessing both of these areas. An evaluative construct is one which may be particularly informative as it is likely that carers who perceive a reduction in their partner's essential goodness will be experiencing a high level of strain. Grid technique may also prove to be useful in the counselling process to provide various foci for exploring with the carer. The implications of these results for the counselling of carers are that the changes
affecting their partner's role fulfilment will be a useful issue to explore with them. The loss of an important source of positive feedback and reward is likely to be central in this.

7) Conclusion

It is clear from the present study that carers experience strain as a result of the various demands of care-giving. The extent of this is such that their mental health is worse than elderly people not caring for a dementing partner. Spouses' Image of their partner was found to have changed and amongst carers this corresponded to the sequelae of dementia. The predicted relationship between Image change and strain was not found. Image change was related to mental health overall but this seems to be due to the implications which such changes have for the extent to which the partner fulfils their role.

Poorer mental health was associated with a lower degree of current role fulfilment in both groups. The degree of role fulfilment was lower amongst carers due to the sort of changes in their partners resulting from dementia. Greater strain in carers was found to be associated with a lower degree of current role fulfilment. In particular, this concerned the extent to which their partner fulfilled the ideal in terms of being caring, open, and reciprocative. It was suggested that these reflect the reduction in an important source of positive self regard.

The measure of previous role fulfilment was found to be a valid indicator of the quality of the premorbid relationship vis a vis how positive and nurturing it was. In general, the quality of the past relationship was no better or worse for carers than others their age. However, psychological well-being was worse amongst carers where the quality of the premorbid relationship was poor in terms of these aspects.
Two other constructs were found to be particularly important concerning carers' psychological well-being. One of these concerned the extent of contentedness which they perceived in their partner. It was suggested that psychological well-being was worse when the partner was perceived as less contented than ideal because carers feel a sense of responsibility for their partner's happiness. Goodness was another important construct, probably reflecting a general evaluative factor. This was suggested to represent a distillation of the implications of the changes in terms of how good their partner is felt to be. Several reasons were suggested why this relates to the carer's psychological well-being. These included guilt, referent self value, and maintenance of balance in the relationship.

The results of the present study have implications for work with carers. They offer some direction for the assessment of the carer's psychological well-being and the possible factors contributing to this. They also indicate some possible foci for exploration in counselling and supporting the carer.

Finally, the various measures of psychological well-being used here were not found to be equivalent. Thus it seems that aspects of the care-giving experience will affect different facets of psychological well-being. For this reason it is important to use more than one measure when studying factors associated with strain in carers.
REFERENCES


### APPENDIX 1

**REPERTORY GRID (I)**

Person being described- ___________ by __________

<table>
<thead>
<tr>
<th>Category</th>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forgetful</td>
<td></td>
<td>Unforgettable</td>
</tr>
<tr>
<td>Honest</td>
<td></td>
<td>Dishonest</td>
</tr>
<tr>
<td>Tough</td>
<td></td>
<td>Tender</td>
</tr>
<tr>
<td>Outgoing</td>
<td></td>
<td>Reserved</td>
</tr>
<tr>
<td>Dominant</td>
<td></td>
<td>Submissive</td>
</tr>
<tr>
<td>Passive</td>
<td></td>
<td>Active</td>
</tr>
<tr>
<td>Confident</td>
<td></td>
<td>Lacking Confidence</td>
</tr>
<tr>
<td>Good</td>
<td></td>
<td>Bad</td>
</tr>
<tr>
<td>Relies on Self</td>
<td></td>
<td>Relies on Others</td>
</tr>
<tr>
<td>Selfish</td>
<td></td>
<td>Generous</td>
</tr>
<tr>
<td>Calm</td>
<td></td>
<td>Excitable</td>
</tr>
<tr>
<td>Helpful</td>
<td></td>
<td>Unhelpful</td>
</tr>
</tbody>
</table>
APPENDIX 2

REPERTORY GRID (II)

<table>
<thead>
<tr>
<th>Person being described- ___________ by ___________</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forgetful</td>
</tr>
<tr>
<td>Open</td>
</tr>
<tr>
<td>Tough</td>
</tr>
<tr>
<td>Outgoing</td>
</tr>
<tr>
<td>Dominant</td>
</tr>
<tr>
<td>Passive</td>
</tr>
<tr>
<td>Confident</td>
</tr>
<tr>
<td>Good</td>
</tr>
<tr>
<td>Relies on Self</td>
</tr>
<tr>
<td>Selfish</td>
</tr>
<tr>
<td>Calm</td>
</tr>
<tr>
<td>Helpful</td>
</tr>
<tr>
<td>Proud</td>
</tr>
<tr>
<td>Patient</td>
</tr>
<tr>
<td>Selfish</td>
</tr>
</tbody>
</table>

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
## APPENDIX 3

**REPERTORY GRID (III)**

<table>
<thead>
<tr>
<th>Person being described- __________</th>
<th>by __________</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forgetful</td>
<td>Remembers</td>
</tr>
<tr>
<td>Open</td>
<td>Secret</td>
</tr>
<tr>
<td>Tough Hearted</td>
<td>Tender Hearted</td>
</tr>
<tr>
<td>Outgoing</td>
<td>Reserved</td>
</tr>
<tr>
<td>Gets their own way</td>
<td>Goes along with others</td>
</tr>
<tr>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Confident</td>
<td>Lacking Confidence</td>
</tr>
<tr>
<td>Good</td>
<td>All right</td>
</tr>
<tr>
<td>Relies on Self</td>
<td>Relies on Others</td>
</tr>
<tr>
<td>Gets ruffled</td>
<td>Doesn’t get ruffled</td>
</tr>
<tr>
<td>Thinks of their own needs</td>
<td>Thinks of other’s needs</td>
</tr>
<tr>
<td>Helpful</td>
<td>Not Helpful</td>
</tr>
<tr>
<td>Patient</td>
<td>Impatient</td>
</tr>
<tr>
<td>Proud</td>
<td>Modest</td>
</tr>
<tr>
<td>Contented</td>
<td>Not contented</td>
</tr>
<tr>
<td>Optimistic</td>
<td>Pessimistic</td>
</tr>
<tr>
<td>Modest</td>
<td>Not Modest</td>
</tr>
<tr>
<td>Gives as much in return as they get from others</td>
<td>Doesn’t give as much in return as they get from others</td>
</tr>
</tbody>
</table>
# APPENDIX 4

## THE GENERAL HEALTH QUESTIONNAIRE

GHQ28

David Goldberg

Please read this carefully.

We should like to know if you have had any medical complaints and how your health has been in general, over the past few weeks. Please answer ALL the question on the following pages simply by underlining the answer which you think most nearly applies to you. Remember that we want to know about present and recent complaints, not those that you had in the past.

It is important that you try to answer ALL the questions. Thank you very much for your cooperation.

<table>
<thead>
<tr>
<th>Have you recently</th>
<th>Better than usual</th>
<th>Same as usual</th>
<th>Worse than usual</th>
<th>Much worse than usual</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 1 -been feeling perfectly well and in good health?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>A 2 -been feeling in need of a good tonic?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>A 3 -been feeling run down and out of sorts?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>A 4 -felt that you are ill?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>A 5 -been getting any pains in your head?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>A 6 -been getting a feeling of tightness or pressure in your head?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>A 7 -been having hot or cold spells?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>B 1 -lost much sleep over worry?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>B 2 -had difficulty in staying asleep once you are off?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>B 3 -felt constantly under strain?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>B 4 -been getting edgy and bad-tempered?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>B 5 -been getting scared or panicky for no good reason?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>B 6 -found everything getting on top of you?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
<tr>
<td>B 7 -been feeling nervous and strung-up all the time?</td>
<td>Not at all</td>
<td>No more than usual</td>
<td>Rather more than usual</td>
<td>Much more than usual</td>
</tr>
</tbody>
</table>
Have you recently

<table>
<thead>
<tr>
<th>C 1</th>
<th>-been managing to keep yourself busy and occupied?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>More so than usual</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>C 2</td>
<td>-been taking longer over the things you do?</td>
</tr>
<tr>
<td></td>
<td>Quicker than usual</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>C 3</td>
<td>-felt on the whole you were doing things well?</td>
</tr>
<tr>
<td></td>
<td>Better than usual</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>C 4</td>
<td>-been satisfied with the way you've carried out your task?</td>
</tr>
<tr>
<td></td>
<td>More satisfied</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>C 5</td>
<td>-felt you were playing a useful part in things?</td>
</tr>
<tr>
<td></td>
<td>More so than usual</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>C 6</td>
<td>-felt capable of making decisions about things?</td>
</tr>
<tr>
<td></td>
<td>More so than usual</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>C 7</td>
<td>-been able to enjoy your normal day-to-day activities?</td>
</tr>
<tr>
<td></td>
<td>More so than usual</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>D 1</td>
<td>-been thinking of yourself as a worthless person?</td>
</tr>
<tr>
<td></td>
<td>Not at all</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>D 2</td>
<td>-felt that life is entirely hopeless?</td>
</tr>
<tr>
<td></td>
<td>Not at all</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>D 3</td>
<td>-felt that life isn't worth living?</td>
</tr>
<tr>
<td></td>
<td>Not at all</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>D 4</td>
<td>-thought of the possibility that you might make away with yourself?</td>
</tr>
<tr>
<td></td>
<td>Definitely not</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>D 5</td>
<td>-found at times you couldn’t do anything because your nerves were too bad?</td>
</tr>
<tr>
<td></td>
<td>Not at all</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>D 6</td>
<td>-found yourself wishing you were dead and away from it all?</td>
</tr>
<tr>
<td></td>
<td>Not at all</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
<tr>
<td>D 7</td>
<td>-found that the idea of taking your own life kept coming into your mind?</td>
</tr>
<tr>
<td></td>
<td>Definitely not</td>
</tr>
<tr>
<td></td>
<td>than usual</td>
</tr>
</tbody>
</table>

© D. Goldberg & The Institute of Psychiatry 1981
APPENDIX 5

STRAIN SCALE (Greene et al 1982)

1) Do you ever feel that you can no longer cope with the situation? 
   Completed by
   Always  Frequently  Sometimes  Rarely  Never
   
2) Do you ever feel that you need a break? 
   Always  Frequently  Sometimes  Rarely  Never
   
3) Do you ever get depressed by the situation? 
   Always  Frequently  Sometimes  Rarely  Never
   
4) Has your own health suffered at all? 
   Always  Frequently  Sometimes  Rarely  Never
   
5) Do you ever worry about accidents happening to _______? 
   Always  Frequently  Sometimes  Rarely  Never
   
6) Do you ever feel that there will be no end to the problem? 
   Always  Frequently  Sometimes  Rarely  Never
   
7) Do you find it difficult to get away on holiday? 
   Considerably  Quite A lot  Moderately  A Little  Not at All
   
8) How much has your social life been affected? 
   Considerably  Quite A lot  Moderately  A Little  Not at All
   
9) How much has the household routine been affected? 
   Considerably  Quite A lot  Moderately  A Little  Not at All
   
10) Is your sleep interrupted by _________? 
    Always  Frequently  Sometimes  Rarely  Never

11) Has your standard of living been affected? 
    Considerably  Quite A lot  Moderately  A Little  Not at All

12) Do you ever feel embarrassed by _________? 
    Always  Frequently  Sometimes  Rarely  Never

13) Are you at all prevented from having visitors? 
    Considerably  Quite A lot  Moderately  A Little  Not at All

14) Do you ever get cross and angry with _________? 
    Always  Frequently  Sometimes  Rarely  Never

15) Do you ever feel frustrated at times with___________? 
    Always  Frequently  Sometimes  Rarely  Never

Personal Distress   (1-6)   ______
Domestic Upset     (7-11)   ______
Negative Feelings  (12-15)   ______
Total RS score     ______
**APPENDIX 6**

**Behaviour And Disturbance Checklist (Greene et al 1982)**

Completed by ___________________ Concerning ___________________

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Always</th>
<th>Frequently</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>take part in family conversations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2)</td>
<td>read newspapers, magazines, etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3)</td>
<td>show an interest in news about friends and relatives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4)</td>
<td>start and maintain a sensible conversation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5)</td>
<td>respond sensibly when spoken to</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6)</td>
<td>understand what is being said to him/her</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7)</td>
<td>watch and follow what is happening on TV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8)</td>
<td>able to keep themself busy doing things</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9)</td>
<td>sit around doing nothing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10)</td>
<td>fail to recognise people</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11)</td>
<td>get mixed up about where he/she is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12)</td>
<td>get mixed up about the day/year/etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13)</td>
<td>have prevented from wandering outside the house</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14)</td>
<td>hoard useless things</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15)</td>
<td>talk nonsense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16)</td>
<td>appear restless and agitated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17)</td>
<td>get lost in the house</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18)</td>
<td>wander outside the house at night</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19)</td>
<td>wander outside the house and get lost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20)</td>
<td>endanger himself/herself</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21)</td>
<td>pace up and down and wring their hands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22)</td>
<td>wander off the subject</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23)</td>
<td>talk out loud to themself</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Behaviour And Disturbance Checklist (Greene et al 1982)

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>seem lost in a world of their own</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Frequently</td>
</tr>
<tr>
<td>25</td>
<td>mood change for no apparent reason</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Frequently</td>
</tr>
<tr>
<td>26</td>
<td>become irritable and easily upset</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Frequently</td>
</tr>
<tr>
<td>27</td>
<td>go on and on about certain things</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Frequently</td>
</tr>
<tr>
<td>28</td>
<td>accuse people of things</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Frequently</td>
</tr>
<tr>
<td>29</td>
<td>become angry and threatening</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Frequently</td>
</tr>
<tr>
<td>30</td>
<td>appear unhappy and depressed</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Frequently</td>
</tr>
<tr>
<td>31</td>
<td>talk all the time</td>
<td>Never</td>
<td>Rarely</td>
<td>Sometimes</td>
<td>Frequently</td>
</tr>
</tbody>
</table>

Apathetic-withdrawn (1-11) _____

Active Disturbed (10-24) _____

'Mood Disturbance (25-31) _____

TOTAL BMD SCORE _____
APPENDIX 7

ISCHEMIA SCORE
(Hachinski et al 1975)
(derived from Mayer-Gross, Slater, & Roth: 3rd ed.)

Patient's Name- _______________ D.o.B.-

Diagnosis- ___________________

Circle One

1) Was the onset of dementia abrupt? Yes / No (2)
2) Have there been fluctuation in the severity of their condition? Yes / No (2)
3) Do they have a history of strokes? Yes / No (2)
4) Are there focal neurological symptoms? Yes / No (2)
5) Are there focal neurological signs? Yes / No (2)
6) Has deterioration occurred in a stepwise fashion? Yes / No (1)
7) Is there marked nocturnal confusion? Yes / No (1)
8) Has there been relative preservation of their premorbid personality? Yes / No (1)
9) Is there now or have there been periods of depression? Yes / No (1)
10) Do they have somatic complaints such as headaches, giddiness, tinnitus, general malaise, or pericardial discomfort? Yes / No (1)
11) Are they emotionally labile with a tendency for sudden emotional expression (i.e. "emotional incontinence")? Yes / No (1)
12) Do they have a history of hypertension? Yes / No (1)
13) Is there evidence of associated arteriosclerosis? Yes / No (1)
### APPENDIX 8

**CLIFTON ASSESSMENT PROCEDURES FOR THE ELDERLY (CAPE)**

**Information & Orientation Scale (I/O Scale)**

<table>
<thead>
<tr>
<th>I/O Scale</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Name</td>
<td></td>
</tr>
<tr>
<td>2 Age</td>
<td></td>
</tr>
<tr>
<td>3 D.o.B.</td>
<td></td>
</tr>
<tr>
<td>4 Ward/Place</td>
<td></td>
</tr>
<tr>
<td>5 Address/Hospital</td>
<td></td>
</tr>
<tr>
<td>6 City/Town</td>
<td></td>
</tr>
<tr>
<td>7 P.M.</td>
<td></td>
</tr>
<tr>
<td>8 U.S. President</td>
<td></td>
</tr>
<tr>
<td>9 Colour of Flag</td>
<td></td>
</tr>
<tr>
<td>10 Day</td>
<td></td>
</tr>
<tr>
<td>11 Month</td>
<td></td>
</tr>
<tr>
<td>12 Year</td>
<td></td>
</tr>
</tbody>
</table>

Total Score: _______

*(max.=12)*
APPENDIX 9

Table of t-Test Results

Results of t-Tests Comparing Dementia & Control Groups on Various Measures and Variables

A) Measures/Variables for which there was no significant difference.

<table>
<thead>
<tr>
<th>Measure</th>
<th>t Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of Carer / Subject</td>
<td>-1.02</td>
<td>(.31)</td>
</tr>
<tr>
<td>Age of Patient / Partner</td>
<td>.30</td>
<td>(.77)</td>
</tr>
<tr>
<td>Length of Marriage</td>
<td>.48</td>
<td>(.63)</td>
</tr>
<tr>
<td>Role Fulfilment Before (RFB)</td>
<td>1.73</td>
<td>(.09)</td>
</tr>
<tr>
<td>Esteem</td>
<td>.84</td>
<td>(.84)</td>
</tr>
</tbody>
</table>

B) Measures/Variables for which there was a significant difference

<table>
<thead>
<tr>
<th>Measure</th>
<th>t Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image Change (IC)</td>
<td>7.96</td>
<td>(p&lt;.001)</td>
</tr>
<tr>
<td>Role Fulfilment Now (RFN)</td>
<td>8.17</td>
<td>(p&lt;.001)</td>
</tr>
<tr>
<td>General Health Questionnaire (GHQ)</td>
<td>6.42</td>
<td>(p&lt;.001)</td>
</tr>
</tbody>
</table>
APPENDIX 10

Correlation Between all Measures in the Combined Group

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>RFN</th>
<th>RFB</th>
<th>Esteem</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFN</td>
<td>.93***</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFB</td>
<td>.18</td>
<td>.41**</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Esteem</td>
<td>-.05</td>
<td>.15</td>
<td>.39**</td>
<td>-</td>
</tr>
<tr>
<td>GHQ</td>
<td>.51***</td>
<td>.63***</td>
<td>.37**</td>
<td>.51***</td>
</tr>
</tbody>
</table>

Correlation Between all Measures Within the Control Group

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>RFN</th>
<th>RFB</th>
<th>Esteem</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFN</td>
<td>.75***</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFB</td>
<td>.46*</td>
<td>.60**</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Esteem</td>
<td>-.09</td>
<td>.22</td>
<td>.24</td>
<td>-</td>
</tr>
<tr>
<td>GHQ</td>
<td>.23</td>
<td>.31</td>
<td>-.11</td>
<td>.27</td>
</tr>
</tbody>
</table>

***p<.001  ** p<.01  *p<.05

IC Total image change  
RFN Total role fulfilment now  
RFB Total role fulfilment before  
CRF Total change in role fulfilment  
GHQ General Health Questionnaire  
B&M Behaviour & Mood Checklist
### APPENDIX 11

Correlations Between all Measures Within the Dementia Group

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>RFN</th>
<th>RFB</th>
<th>Esteem</th>
<th>GHQ</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFN</td>
<td>.82**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFB</td>
<td>-.26</td>
<td>.24</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esteem</td>
<td>-.02</td>
<td>.34</td>
<td>.57**</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GHQ</td>
<td>.26</td>
<td>.53**</td>
<td>.42*</td>
<td>.80***</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Strain</td>
<td>.26</td>
<td>.39*</td>
<td>.19</td>
<td>.39*</td>
<td>.65***</td>
<td>-</td>
</tr>
<tr>
<td>B&amp;M</td>
<td>.46*</td>
<td>.62**</td>
<td>.10</td>
<td>.15</td>
<td>.32</td>
<td>.38*</td>
</tr>
</tbody>
</table>

** ***p<.001  ** p<.01  * p<.05

- **IC**: Total image change
- **RFN**: Total role fulfilment now
- **RFB**: Total role fulfilment before
- **GHQ**: General Health Questionnaire
- **B&M**: Behaviour & Mood Checklist
### APPENDIX 12

**Rank Order of Mean IC Scores of All Constructs in Both Groups**

<table>
<thead>
<tr>
<th>Score</th>
<th>Rank</th>
<th>Mean IC Score</th>
<th>Construct</th>
<th>Construct</th>
<th>Mean IC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.45</td>
<td>1</td>
<td>4.45</td>
<td>C Memory</td>
<td>Dominance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.35</td>
<td>C Self Reliance</td>
<td>Calmness</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4.20</td>
<td>C Activeness</td>
<td>Memory</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3.95</td>
<td>C Reciprocation</td>
<td>Reservedness</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>3.90</td>
<td>C Confidence</td>
<td>Confidence</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>3.60</td>
<td>C Selfishness</td>
<td>Self reliance</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>3.35</td>
<td>C Helpfulness</td>
<td>Selfishness</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>3.15</td>
<td>H Optimism</td>
<td>Patience</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>3.10</td>
<td>H Secrecy</td>
<td>Active ness</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>2.55</td>
<td>H Caring</td>
<td>Optimism</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.55</td>
<td>Dominance</td>
<td>Modesty</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>2.40</td>
<td>Reservedness</td>
<td>Secrecy</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>2.10</td>
<td>H Goodness</td>
<td>Helpfulness</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>2.05</td>
<td>C Patience</td>
<td>Goodness</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>1.80</td>
<td>H Modesty</td>
<td>Reciprocation</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>1.40</td>
<td>C Calmness</td>
<td>Caring</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>0.75</td>
<td>C Contentedness</td>
<td>Contentedness</td>
<td>0.25</td>
</tr>
</tbody>
</table>

There is no significant correlation between the two sets of rank ordered mean IC scores ($r = .17$ corrected for tied ranks).

"C" indicates constructs which were predicted to change the most.

"H" indicates constructs predicted to be stable, i.e. show little change.
APPENDIX 13

Correlations Between Separate IC Constructs and Various Measures Within the Dementia Group

Correlations concerning the size of IC score

<table>
<thead>
<tr>
<th>Construct</th>
<th>Strain</th>
<th>GHQ</th>
<th>Esteem</th>
<th>B&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>.10</td>
<td>-.10</td>
<td>-.10</td>
<td>.35</td>
</tr>
<tr>
<td>Secrecy</td>
<td>.12</td>
<td>.18</td>
<td>.05</td>
<td>.17</td>
</tr>
<tr>
<td>Caring</td>
<td>.40*</td>
<td>.35</td>
<td>.10</td>
<td>.61**</td>
</tr>
<tr>
<td>Reservedness</td>
<td>-.04</td>
<td>-.21</td>
<td>-.28</td>
<td>.02</td>
</tr>
<tr>
<td>Dominance</td>
<td>.16</td>
<td>.11</td>
<td>-.10</td>
<td>.42</td>
</tr>
<tr>
<td>Activity</td>
<td>-.04</td>
<td>.02</td>
<td>-.08</td>
<td>.07</td>
</tr>
<tr>
<td>Confidence</td>
<td>.23</td>
<td>.19</td>
<td>-.04</td>
<td>.11</td>
</tr>
<tr>
<td>Goodness</td>
<td>.55**</td>
<td>.531*</td>
<td>.43*</td>
<td>.32</td>
</tr>
<tr>
<td>Self Reliance</td>
<td>-.11</td>
<td>.16</td>
<td>.06</td>
<td>-.32</td>
</tr>
<tr>
<td>Calmness</td>
<td>-.27</td>
<td>-.23</td>
<td>-.31</td>
<td>-.02</td>
</tr>
<tr>
<td>Selfishness</td>
<td>-.01</td>
<td>-.01</td>
<td>.04</td>
<td>.38</td>
</tr>
<tr>
<td>Helpfulness</td>
<td>.20</td>
<td>.24</td>
<td>.03</td>
<td>.52**</td>
</tr>
<tr>
<td>Patience</td>
<td>-.24</td>
<td>.06</td>
<td>-.20</td>
<td>.24</td>
</tr>
<tr>
<td>Contentedness</td>
<td>.49**</td>
<td>.17</td>
<td>.06</td>
<td>.26</td>
</tr>
<tr>
<td>Optimism</td>
<td>.19</td>
<td>.22</td>
<td>.08</td>
<td>.16</td>
</tr>
<tr>
<td>Modesty</td>
<td>.24</td>
<td>.24</td>
<td>.05</td>
<td>.30</td>
</tr>
<tr>
<td>Reciprocation</td>
<td>.21</td>
<td>.13</td>
<td>-.14</td>
<td>.22</td>
</tr>
</tbody>
</table>

** p<.01     * p<.05
## APPENDIX 13 (cont.)

**Correlations concerning the direction of IC difference.**

<table>
<thead>
<tr>
<th>Construct</th>
<th>Strain</th>
<th>GHQ</th>
<th>Esteem</th>
<th>B&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>-.10</td>
<td>.10</td>
<td>-.05</td>
<td>-.38</td>
</tr>
<tr>
<td>Secrecy</td>
<td>.37</td>
<td>.36</td>
<td>.23</td>
<td>.17</td>
</tr>
<tr>
<td>Caring</td>
<td>-.40</td>
<td>.53*</td>
<td>.41</td>
<td>-.39</td>
</tr>
<tr>
<td>Reservedness</td>
<td>.24</td>
<td>-.08</td>
<td>.06</td>
<td>.12</td>
</tr>
<tr>
<td>Dominance</td>
<td>-.02</td>
<td>.10</td>
<td>.08</td>
<td>-.27</td>
</tr>
<tr>
<td>Activity</td>
<td>.04</td>
<td>.02</td>
<td>.14</td>
<td>.05</td>
</tr>
<tr>
<td>Confidence</td>
<td>.30</td>
<td>.24</td>
<td>-.02</td>
<td>.06</td>
</tr>
<tr>
<td>Goodness</td>
<td>.17</td>
<td>-.03</td>
<td>.08</td>
<td>.26</td>
</tr>
<tr>
<td>Self Reliance</td>
<td>.01</td>
<td>.15</td>
<td>.03</td>
<td>.08</td>
</tr>
<tr>
<td>Calmness</td>
<td>.16</td>
<td>.18</td>
<td>.21</td>
<td>-.01</td>
</tr>
<tr>
<td>Selfishness</td>
<td>-.15</td>
<td>-.15</td>
<td>-.05</td>
<td>-.44</td>
</tr>
<tr>
<td>Helpfulness</td>
<td>-.02</td>
<td>.10</td>
<td>-.22</td>
<td>.32</td>
</tr>
<tr>
<td>Patience</td>
<td>-.19</td>
<td>.11</td>
<td>-.23</td>
<td>.18</td>
</tr>
<tr>
<td>Contentedness</td>
<td>.40</td>
<td>.17</td>
<td>.07</td>
<td>.23</td>
</tr>
<tr>
<td>Optimism</td>
<td>.15</td>
<td>.12</td>
<td>-.39</td>
<td>-.01</td>
</tr>
<tr>
<td>Modesty</td>
<td>.60*</td>
<td>.61**</td>
<td>.05</td>
<td>.48</td>
</tr>
<tr>
<td>Reciprocation</td>
<td>.39</td>
<td>.15</td>
<td>-.28</td>
<td>.42</td>
</tr>
</tbody>
</table>

**NB** In calculating correlations concerning the direction of discrepancy, cases in which the carer felt the construct was not applicable (when rating 'Now') were excluded from the analysis therefore the size of N varied. This accounts for smaller correlations being significant in some cases when larger correlations were not.
APPENDIX 14

Correlations Between Separate RFB Constructs and Various Measures Within the Dementia Group

Correlations concerning the size of RFB score

<table>
<thead>
<tr>
<th>Construct</th>
<th>Strain</th>
<th>GHQ</th>
<th>Esteem</th>
<th>B&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>.36</td>
<td>.53*</td>
<td>.54**</td>
<td>.15</td>
</tr>
<tr>
<td>Secrecy</td>
<td>-.27</td>
<td>-.02</td>
<td>.23</td>
<td>-.20</td>
</tr>
<tr>
<td>Caring</td>
<td>-.40*</td>
<td>-.16</td>
<td>-.17</td>
<td>-.17</td>
</tr>
<tr>
<td>Reservedness</td>
<td>.14</td>
<td>-.01</td>
<td>.13</td>
<td>.05</td>
</tr>
<tr>
<td>Dominance</td>
<td>.31</td>
<td>.33</td>
<td>.17</td>
<td>.12</td>
</tr>
<tr>
<td>Activity</td>
<td>.14</td>
<td>.17</td>
<td>.23</td>
<td>.25</td>
</tr>
<tr>
<td>Confidence</td>
<td>.02</td>
<td>.19</td>
<td>.37</td>
<td>.33</td>
</tr>
<tr>
<td>Goodness</td>
<td>.40*</td>
<td>.13</td>
<td>.08</td>
<td>.06</td>
</tr>
<tr>
<td>Self Reliance</td>
<td>-.15</td>
<td>-.25</td>
<td>-.06</td>
<td>-.04</td>
</tr>
<tr>
<td>Calmness</td>
<td>.25</td>
<td>.46*</td>
<td>.55**</td>
<td>.21</td>
</tr>
<tr>
<td>Selfishness</td>
<td>-.14</td>
<td>.04</td>
<td>-.11</td>
<td>.45*</td>
</tr>
<tr>
<td>Helpfulness</td>
<td>.42*</td>
<td>.34</td>
<td>.27</td>
<td>.40*</td>
</tr>
<tr>
<td>Patience</td>
<td>.19</td>
<td>.40*</td>
<td>.58**</td>
<td>.12</td>
</tr>
<tr>
<td>Contentedness</td>
<td>.57**</td>
<td>.79**</td>
<td>.63***</td>
<td>.46*</td>
</tr>
<tr>
<td>Optimism</td>
<td>.18</td>
<td>.34</td>
<td>.55**</td>
<td>.22</td>
</tr>
<tr>
<td>Modesty</td>
<td>-.29</td>
<td>-.11</td>
<td>-.21</td>
<td>-.31</td>
</tr>
<tr>
<td>Reciprocation</td>
<td>-.11</td>
<td>.03</td>
<td>.16</td>
<td>-.17</td>
</tr>
</tbody>
</table>

Correlations concerning the direction of RFB discrepancy

<table>
<thead>
<tr>
<th>Construct</th>
<th>Strain</th>
<th>GHQ</th>
<th>Esteem</th>
<th>B&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>.26</td>
<td>.46</td>
<td>.49</td>
<td>.15</td>
</tr>
<tr>
<td>Secrecy</td>
<td>-.24</td>
<td>-.23</td>
<td>-.37</td>
<td>-.32</td>
</tr>
<tr>
<td>Caring</td>
<td>.21</td>
<td>.21</td>
<td>.04</td>
<td>.15</td>
</tr>
<tr>
<td>Reservedness</td>
<td>.47*</td>
<td>.27</td>
<td>-.09</td>
<td>.01</td>
</tr>
<tr>
<td>Dominance</td>
<td>.13</td>
<td>.25</td>
<td>.30</td>
<td>.06</td>
</tr>
<tr>
<td>Activity</td>
<td>.03</td>
<td>.01</td>
<td>.04</td>
<td>.06</td>
</tr>
<tr>
<td>Confidence</td>
<td>-.03</td>
<td>-.02</td>
<td>-.03</td>
<td>-.16</td>
</tr>
<tr>
<td>Goodness</td>
<td>-.20</td>
<td>-.31</td>
<td>-.16</td>
<td>.01</td>
</tr>
<tr>
<td>Self Reliance</td>
<td>.06</td>
<td>.08</td>
<td>.03</td>
<td>-.07</td>
</tr>
<tr>
<td>Calmness</td>
<td>.40</td>
<td>.50*</td>
<td>.25</td>
<td>.31</td>
</tr>
<tr>
<td>Selfishness</td>
<td>.05</td>
<td>.31</td>
<td>.10</td>
<td>.10</td>
</tr>
<tr>
<td>Helpfulness</td>
<td>-.21</td>
<td>-.16</td>
<td>-.23</td>
<td>-.35</td>
</tr>
<tr>
<td>Patience</td>
<td>-.06</td>
<td>-.26</td>
<td>-.45*</td>
<td>-.10</td>
</tr>
<tr>
<td>Contentedness</td>
<td>-.27</td>
<td>-.65**</td>
<td>-.40</td>
<td>-.21</td>
</tr>
<tr>
<td>Optimism</td>
<td>-.28</td>
<td>-.37</td>
<td>-.48*</td>
<td>-.38</td>
</tr>
<tr>
<td>Modesty</td>
<td>.05</td>
<td>.06</td>
<td>-.20</td>
<td>-.05</td>
</tr>
<tr>
<td>Reciprocation</td>
<td>-.04</td>
<td>-.32</td>
<td>-.31</td>
<td>.31</td>
</tr>
</tbody>
</table>

*** p<.001    ** p<.01     * p<.05
# APPENDIX 15

## Correlations Between Separate RPN Constructs and Various Measures in the Dementia Group

### Correlations concerning the size of RFN score

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Strain</th>
<th>GHQ</th>
<th>Esteem</th>
<th>B&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>.37</td>
<td>.33</td>
<td>.44*</td>
<td>.60**</td>
</tr>
<tr>
<td>Secrecy</td>
<td>.30</td>
<td>.35</td>
<td>.26</td>
<td>.37</td>
</tr>
<tr>
<td>Caring</td>
<td>.28</td>
<td>.29</td>
<td>.18</td>
<td>.42</td>
</tr>
<tr>
<td>Reservedness</td>
<td>-.34</td>
<td>-.22</td>
<td>-.11</td>
<td>.08</td>
</tr>
<tr>
<td>Dominance</td>
<td>.08</td>
<td>.14</td>
<td>.25</td>
<td>.30</td>
</tr>
<tr>
<td>Activity</td>
<td>.05</td>
<td>.14</td>
<td>-.06</td>
<td>.10</td>
</tr>
<tr>
<td>Confidence</td>
<td>.26</td>
<td>.23</td>
<td>.08</td>
<td>.27</td>
</tr>
<tr>
<td>Goodness</td>
<td>.37</td>
<td>.40*</td>
<td>.42*</td>
<td>.30</td>
</tr>
<tr>
<td>Self Reliance</td>
<td>.53</td>
<td>.35</td>
<td>.22</td>
<td>.40</td>
</tr>
<tr>
<td>Calmness</td>
<td>.27</td>
<td>.33</td>
<td>.17</td>
<td>.31</td>
</tr>
<tr>
<td>Selfishness</td>
<td>.08</td>
<td>.12</td>
<td>.14</td>
<td>.02</td>
</tr>
<tr>
<td>Helpfulness</td>
<td>.14</td>
<td>.17</td>
<td>-.04</td>
<td>.40</td>
</tr>
<tr>
<td>Patience</td>
<td>-.19</td>
<td>.24</td>
<td>.20</td>
<td>.29</td>
</tr>
<tr>
<td>Contentedness</td>
<td>.55**</td>
<td>.64***</td>
<td>.41</td>
<td>.40</td>
</tr>
<tr>
<td>Optimism</td>
<td>.23</td>
<td>.25</td>
<td>.21</td>
<td>.35</td>
</tr>
<tr>
<td>Modesty</td>
<td>.24</td>
<td>.20</td>
<td>.08</td>
<td>.24</td>
</tr>
<tr>
<td>Reciprocation</td>
<td>.53**</td>
<td>.42</td>
<td>.10</td>
<td>.36</td>
</tr>
</tbody>
</table>

*"**" p<.001  ** p<.01  *" p<.05
Correlations concerning the direction of RFN discrepancy

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Strain</th>
<th>GHO</th>
<th>Esteem</th>
<th>B&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memoev</td>
<td>.37</td>
<td>.33</td>
<td>.53*</td>
<td>.59**</td>
</tr>
<tr>
<td>Secrecv</td>
<td>-.63**</td>
<td>.56*</td>
<td>-.47</td>
<td>-.46</td>
</tr>
<tr>
<td>Caring</td>
<td>.57*</td>
<td>.70*</td>
<td>.51*</td>
<td>.44*</td>
</tr>
<tr>
<td>Reservedness</td>
<td>.30</td>
<td>.33</td>
<td>-.18</td>
<td>-.05</td>
</tr>
<tr>
<td>Dominance</td>
<td>-.01</td>
<td>-.12</td>
<td>.02</td>
<td>.27</td>
</tr>
<tr>
<td>Activtv</td>
<td>-.03</td>
<td>-.02</td>
<td>-.15</td>
<td>-.04</td>
</tr>
<tr>
<td>Confidence</td>
<td>-.32</td>
<td>-.26</td>
<td>-.06</td>
<td>-.20</td>
</tr>
<tr>
<td>Goodness</td>
<td>-.33</td>
<td>-.29</td>
<td>-.20</td>
<td>-.25</td>
</tr>
<tr>
<td>Self Reliance</td>
<td>.05</td>
<td>-.29</td>
<td>-.08</td>
<td>-.34</td>
</tr>
<tr>
<td>Calmness</td>
<td>.22</td>
<td>.30</td>
<td>.06</td>
<td>.27</td>
</tr>
<tr>
<td>Selfishness</td>
<td>.25</td>
<td>.36</td>
<td>.24</td>
<td>.40</td>
</tr>
<tr>
<td>HELfulness</td>
<td>.02</td>
<td>-.10</td>
<td>.19</td>
<td>-.37</td>
</tr>
<tr>
<td>Patience</td>
<td>.11</td>
<td>-.27</td>
<td>-.11</td>
<td>-.22</td>
</tr>
<tr>
<td>Contentedness</td>
<td>-.45</td>
<td>-.58**</td>
<td>-.33</td>
<td>-.30</td>
</tr>
<tr>
<td>Optimism</td>
<td>-.20</td>
<td>-.21</td>
<td>-.01</td>
<td>-.26</td>
</tr>
<tr>
<td>Modestv</td>
<td>-.46</td>
<td>-.59*</td>
<td>-.11</td>
<td>-.50*</td>
</tr>
<tr>
<td>Reciprocation</td>
<td>-.46*</td>
<td>-.40</td>
<td>.16</td>
<td>-.30</td>
</tr>
</tbody>
</table>

NB In calculating correlations concerning the direction of discrepancy, cases in which the carer felt the construct was not applicable (when rating 'Now') were excluded from the analysis therefore the size of N varied. This accounts for smaller correlations being significant in some cases when larger correlations were not.
## APPENDIX 16

### Rank Order of Mean RFB Scores for All Constructs in Both Groups

<table>
<thead>
<tr>
<th>Score</th>
<th>Rank</th>
<th>Dementia Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean RFB Score</td>
<td>Construct</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2.40</td>
<td>Reservedness</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1.75</td>
<td>Caring</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1.65</td>
<td>Self Reliance</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1.60</td>
<td>Dominance</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>1.55</td>
<td>Calmness</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>1.55</td>
<td>Selfishness</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>1.30</td>
<td>Optimism</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>1.20</td>
<td>Patience</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>1.15</td>
<td>Goodness</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>1.10</td>
<td>Secrecy</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>0.95</td>
<td>Memory</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>0.90</td>
<td>Reciprocation</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>0.80</td>
<td>Activeness</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>0.80</td>
<td>Confidence</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>0.70</td>
<td>Contentedness</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>0.60</td>
<td>Modesty</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>0.55</td>
<td>Helpfulness</td>
</tr>
</tbody>
</table>

There is a significant correlation between the two sets of rank order mean RFB scores ($r = .67$, corrected for tied ranks).
### APPENDIX 17

**Rank Order of Mean RFN Scores of All Constructs in Both Groups**

| Rank | Mean RFN Score | Construct        | Dementia Group | | | Construct | Control Group | Mean RFN Score |
|------|----------------|------------------|----------------|------------------|----------------|----------------|----------------|
| 1    | 4.70           | Memory           | Dominance      | 2.50             | Patience       | 2.15           |
| 2    | 4.70           | Self Reliance    | Patience       | 2.15             | Optimism       | 1.65           |
| 3    | 4.40           | Selfishness      | Optimism       | 1.65             | Memory         | 1.50           |
| 4    | 4.35           | Confidence       | Memory         | 1.50             | Reservedness   | 1.40           |
| 5    | 4.00           | Reciprocation    | Reservedness   | 1.40             | Calmness       | 1.40           |
| 6    | 3.60           | Activeness       | Calmness       | 1.40             | Goodness       | 1.10           |
| 7    | 3.55           | Optimism         | Goodness       | 1.10             | Secrecy        | 1.05           |
| 8    | 3.30           | Helpfulness      | Secrecy        | 1.05             | Modesty        | 1.05           |
| 9    | 3.05           | Patience         | Modesty        | 1.05             | Reliance       | 1.00           |
| 10   | 3.00           | Reservedness     | Reliance       | 1.00             | Confidence     | 0.95           |
| 11   | 2.95           | Caring           | Confidence     | 0.95             | Activeness     | 0.90           |
| 12   | 2.85           | Secrecy          | Activeness     | 0.90             | Dominance      | 0.85           |
| 13   | 2.55           | Dominance        | Selfishness    | 0.85             | Goodness       | 0.60           |
| 14   | 2.10           | Goodness         | Caring         | 0.60             | Calmness       | 0.55           |
| 15   | 2.05           | Calmness         | Reciprocation  | 0.55             | Helpfulness    | 0.45           |
| 16   | 1.80           | Modesty          | Helpfulness    | 0.45             | Contentedness  | 0.15           |
| 17   | 1.15           | Contentedness    | Contentedness  | 0.15             |                |                |

There is no significant correlation between the two sets of rank order RFN mean scores ($r=0.06$, corrected for tied ranks).